# ABSTRACT

Title of dissertation: UNDERSTANDING THE MOTILITY OF MOLECULAR MOTORS USING THEORY AND SIMULATIONS

Yonathan Goldtzvik, Doctor of Philosophy, 2017

Directed by: Professor Devarajan Thirumalai
Biophysics Program
Institute for Physical Science and Technology,
Department of Chemistry and Biochemistry

Molecular motors are indispensable machines that are in charge of transporting cargoes within living cells. Despite recent advances in the study of these molecules, there is much that we still do not understand regarding the underlying mechanisms that allow them to efficiently move cargoes along polar cellular filaments. In this thesis, I report my investigation on two motor proteins superfamilies, dyneins and kinesins. Using theoretical modeling, we provide fundamental insight into their function.

Dynein is a large motor that transports cargo along microtubules towards their negative pole. Unlike other motors, such as conventional kinesin, the motility of dynein is highly stochastic. We developed a novel theoretical approach, which reproduces a wide variety of its properties, including the unique step size distribution observed in experiments. Furthermore, our model enables us to derive several

simple expressions that can be fitted to experiment, thus providing a physical interpretation.

A less understood aspect of dynein is the complex set of allosteric transitions in response to ATP binding and hydrolysis, and microtubule binding. The resulting conformational transitions propel the motor forward to the minus end of the microtubule. Furthermore, its activity is regulated by external strain. Using coarse grained Brownian dynamics simulations, we show that a couple of insert loops in the AAA2, a sub domain in the AAA+ ring in the motor domain, play an important role in several of the alllosteric pathways.

Kinesins are highly processive motor proteins that transport cargo along microtubules toward their positive poles. Experiments show that the kinesin motor domains propel the motor forward by passing each other in a hand-over-hand motion. However, there is a debate as to whether the motor domains do so in a symmetrical manner or an asymmetrical motion. Using coarse grained Brownian dynamics simulations of the kinesin dimer, we show that the kinesin stepping mechanism is influenced by the size of its cargo. Furthermore, we find that stepping occurs by a combinations of both the symmetric and asymmetric mechanisms. The results I present in this thesis are a testimony that theoretical approaches are invaluable to the study of molecular motors.

# UNDERSTANDING THE MOTILITY OF MOLECULAR MOTORS USING THEORY AND SIMULATIONS

by

Yonathan Goldtzvik

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor Devarajan Thirumalai, Chair
Professor Garegin A. Papoian, Co-Chair
Professor Dorothy Beckett
Professor Jeffery Klauda
Professor George Lorimer
Professor John D. Weeks

# Dedication

To my grandparents, I am who I am because of you.

# Acknowledgments

First I would like to thank my advisor, Prof. Dave Thirumalai for giving me the opportunity to join his research group and to pursue my intellectual curiosities. Over the last five years, Dave has constantly pushed me to be a better scientist and I have benefited greatly from his insight. He created a nurturing research environment in which free discussion and collaboration are encouraged and for that I am grateful.

I would like to thank my defense committee members. Professors Lorimer and Papoian have been on my committee throughout the Ph.D. process and I have always benefited from their advice and insight. I have also learned a great deal from Prof. Weeks, and appreciate him, along with professors Beckett and Klauda, for making the time to join my defense committee.

Special thanks to Prof. Mike Hinczewski and Dr. Mauro Mugnai. During his time as a postdoctoral researcher in the Thirumalai group, Mike was a mentor to me in the truest sense of the word. His door was always open, and he pushed me to tackle new approaches to my research questions. Mauro is both a terrific colleague and mentor. His mathematical prowess and kindness are matched by few and are a source of constant inspiration.

I am grateful for all of the members of the Thirumalai research group with whom I enjoyed many discussions, both scientific and political. In particular, Huong Vu joined the research group at the same time as I did and has been a great colleague. Also, I am grateful to Dr. Shaon Chakrabarti who is always cheerful and made the Ph.D. process seem easy.

I would like to add special thanks to Prof. Robert Benny Gerber and Dr. Moshe Goldstein who taught me how to be a scientist and without whom none of this would be possible.

Thanks to my Washington DC community of friends, who are like a family to me and who always remind me that there is more to life than science.

I cannot thank enough my mother- and father-in-law, Debbie and Eric Goldberg, who created a home away from home for me over these last few years. Their unflagging confidence and support has been invaluable.

I am grateful beyond measure to my parents, Itamar Cwik and Nora Benderski, who never tired of hearing about my endeavors and found my research just as interesting as I did. They have been an endless source of support and I could not have done this without them.

Finally and most importantly, I would like to thank my wife, Sharon Goldtzvik, to whom I believe this thesis belongs just as much as it belongs to me. She has been a constant source of inspiration and always reminded me that we are each the author of our own story. I dedicate this work with love to her.

# Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| CC | Coiled Coil |
| CG | Coarse Grained |
| EM | Electron Microscopy |
| FENE | Finite Extensible Non-Linear Elastic |
| FIONA | Fluorescence Imaging at One Nanometer Accuracy |
| HI | Hydrodynamic Interactions |
| IL | Insert Loop |
| LH | Leading Head |
| LN | Linker |
| MD | Motor Domain |
| MT | Microtubule |
| MTBD | Microtubule Binding Domain |
| MU | Motor Unit |
| MV | Myosin V |
| NL | Neck Linker |
| NM | Non-Stepping Motor |
| PDB | Protein Data Bank |
| Pi | Inorganic Phosphate |
| SM | Stepping Motor |
| SOP | Self Organized Polymer |
| TBS | Target Binding Site |
| TH | Trailing Head |
| ZT | Zhang and Thirumalai |

# Chapter 1

# Introduction

Living cells are faced with the challenge of being able to both regulate their internal environment and affect their surroundings. A crucial element in overcoming these challenges is the ability of cells to convert chemical energy into mechanical work. Motor proteins are the cellular components that enable cells to do so. These molecular machines are involved in a wide variety of processes including cell motility, mechanosensation and cargo transportation along intracellular filaments[1,2,3]. Since these processes are of great importance in the proper functioning of cells, there is a high level of interest in understanding the underlying functional mechanisms of these motors.

There are three major superfamilies of motor proteins - myosins, dyneins and kinesins, all of which share certain structural and functional properties while differing in others[4]. In general, the structure of a typical motor protein consists of a pair of motor domains (MD) that "walks" along cellular filaments by taking discrete steps[5]. Different superfamilies associate with different types of filaments. Myosins typically walk along actin filaments while kinesins and dyneins walk along microtubules (MT)[6,7]. All these filaments have an intrinsic structural directionality or polarity. This allows molecular motors to move preferentially in a particular di-

rection along the filament. While the particular walking mechanism differs from motor to motor, all of these motors share the ability to hydrolyze ATP molecules and couple the hydrolysis to a structural change during the chemomechanical cycle. This conformational transition is often referred to as the "power stroke" and is a key element in the mechanism that allows a motor protein to control the direction of its motion[8,9].

Though crucial, the power stroke in and by itself is not sufficient to propel thee motor along the polar track. The two MDs have to be able to coordinate their stepping to some extent. By coordination we mean that a MD's process of "deciding" to initiate a step depends to some extent on the state of the second MD in the dimer. This coordination is necessary for two reasons. The first is to avoid consecutive stepping of the same MD which would lead to a "stomping" motion[10]. The second reason is to maintain high processivity (the number of consecutive steps a motor can take before detachment from the track) while moving along the filament at a high speed. One of the main factors that determine the motor's speed is the amount of time that each MD waits before stepping again. If this dwell time is short and the MDs do not coordinate their steps, the motor will move with a high speed. However, it will also increase the probability that after one of the MDs begins its step, the second domain will step as well, leading to complete detachment from the filament and termination of the motor's walk. Coordination between the two MDs reduces the likelihood of such simultaneous stepping, thus increasing processivity.

In the last several decades major advancements in single molecule detection methods such as Fluorescence Imaging at One Nanometer Accuracy (FIONA), quantum dots and optical tweezers, allowed scientists to track the motion of individual motor proteins[11,12,13,14]. Experimentalists are able to measure such properties as motor velocities, step sizes, and run length. At the same time, much progress was made in the elucidation of the structural features of motor proteins using X-ray

crystallography and Electron Microscopy (EM) as well as their biochemical properties[9,15,16,17,18,19,20,21]. As the amount of information about molecular motors increases, a more cohesive picture of how these machines work has emerged. However, there is still much we do not understand about the underlying mechanisms of their motion. Our goal in this work is to use the body of available biophysical data to better understand how motor proteins function while focusing on the MT associated superfamilies - kinesins and dyneins.

Experimental methods are physically limited in terms of their spatial and temporal resolution. Furthermore, especially in the case of complex systems, it can be difficult to put all the information together in a way that makes sense. Therefore, we resort to theoretical and computational methods to fill in the gaps in our understanding of molecular motors. Such models can be useful on two levels, in the context of motor proteins. One is the "microscopic" level that deals with structural dynamics such as allosteric transition and conformational changes of the protein. The other is the "macroscopic" level that deals with the structure and function of the motor as a whole and is concerned with its "macroscopic" properties. In this thesis, we show how both approaches can be successfully applied to motor proteins and yield important insights into their working mechanisms.

# Chapter 2

# Dynein: Overview

Dyneins, like kinesins, are motor proteins that move along MTs. However, there seem to be more differences than similarities between the two superfamilies. While in most cases kinesins move predominantly towards the MT's positive end, dyneins are in charge of most of the cellular transportation towards the negative ends of MTs in eukaryotic cells[22]. Furthermore, the structure and stepping mechanisms appears to be quite different when comparing the two superfamilies.

The dynein superfamily can be divided into three major categories - axonemal dyneins, intra-flagellar transport dyneins, and cytoplasmic dyneins[3]. Axonemal dyneins are embedded into the axoneme of cilia and flagella. Their role is to power the motion of these cellular appendages by sliding the MTs relative to each other within the axoneme[23]. Intra-flagellar transport dyneins are involved in the transportation of structural components along the cilia and flagella towards the body of the cell[24,25]. Cytoplasmic dyneins participate in a variety of functions within the cell's body. These include transportation of different organelles such as endosomes, lysosomes, and mitochondria as well as different cargos such as transcription factors and structural components from the cell's periphery towards it nucleus[26,27,28,29,30]. Furthermore, cytoplasmic dyneins can exert tension on different parts of the cells

and are crucial for such processes as positioning of the spindle during cell division[31,32]. Cytoplasmic dynein is the focus of our research and we refer to it simply as dynein from this point onwards.

Perhaps the most obvious reason for the existence of a motor such as dynein is the need for bidirectional transport along MTs. Even more interesting is the fact that the same cargo can bind simultaneously to both dyneins and kinesins. The pulling of a single cargo by multiple motors in opposite directions is well documented and is commonly referred to as a "tug of war"[33,34,35,36]. Furthermore, a growing body of evidence suggests that inhibition of one motor family leads to diminished activity of its counterpart when both motor types are bound to the same cargo[37]. Therefore, the existence of dynein is necessary, not only for transportation of cargo in the negative direction along the MT, but also for regulation of cargo transportation by kinesin in the positive direction.

While kinesins and myosins have structural features in common with G proteins, dynein is distinct and belongs to the AAA+ family of proteins[38]. As such, its structure is quite different, both in fold and size, when compared to the structures of kinesin and myosin. The basic structural element responsible for the dynein function is the MD; however, a dimer of MDs is the minimal construct for dynein to move processively along the MT[39].

The MD itself can be divided into several structural domains: AAA+ ring, linker (LN), stalk, Microtubule Binding Domain (MTBD), and strut (Fig. 2.1)[17,18]. The AAA+ ring consists of six AAA units and is responsible for the conversion of ATP hydrolysis into mechanical work. In most molecular machines of the AAA family such as GroEL, the AAA ring consists of several identical AAA units. That is not the case for dynein. Each of the AAA units in dynein, which are divided into a large and a small subunit, is different and serves different function. The AAA1 unit has been shown to be the main ATP hydrolysis site of the motor[40]. ATP binds

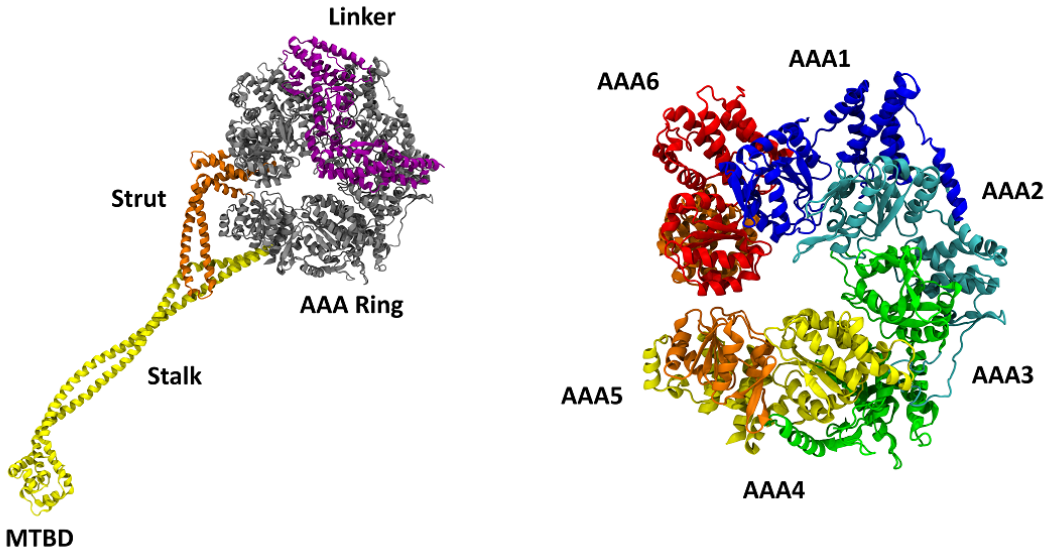


Figure 2.1: Structure of the dynein MD. Left: The MD divided into its components: AAA+ ring (gray), linker (purple), strut (orange), stalk and Microtubule Binding Domain (MTBD) (yellow). Right: The AAA+ ring, divided into the six AAA domains: AAA1 (blue), AAA2 (cyan), AAA3 (green), AAA4 (yellow), AAA5 (orange), and AAA6 (red).

to the cleft between the AAA1 and AAA2 units with high affinity and induces a closure of the cleft[19]. This conformational transition is at the center of the allosteric transitions in the motor and has been shown to control the affinity of the motor for MTs[41]. The AAA2 unit can bind nucleotides as well but does not seem to hydrolyze ATP and in that sense its role appears to be mostly structural. Nevertheless, the AAA2 unit contains two Insert Loops (IL) that have been shown to interact with the LN domain and seem to play an important role in the regulation of the motor's activity[42,43]. The AAA3 unit binds and hydrolyzes ATP, but at a significantly lower rate in comparison to the AAA1 unit[44]. The full purpose of the AAA3 unit is not well understood. However, several studies indicate that it plays an important role in the regulation of the affinity of the motor to the MT[40,44,45]. The AAA4 unit is also capable of binding and hydrolyzing ATP but whether it serves a particular purpose is unclear. The main role of the unit is to serve as the base of the stalk domain that

emerges from the AAA ring. Unlike the AAA units mentioned up till this point, both AAA5 and AAA6 units can neither bind nor hydrolyze ATP. The two units do undergo significant conformational transitions depending on the nucleotide state of the AAA1 unit[17,18,19]. Furthermore, the AAA5 unit serves as the base for the strut domain that connects it to the stalk.

The LN is a large structure at the N-terminal of the MD. It is attached to the AAA1 unit at its base and spans the diameter of the AAA+ ring. The LN serves two major purposes. The first is to connect the MD to a tail domain. The tail domain is responsible in turn for the dimerization of the two MDs as well as attachment to the appropriate cargo[18]. The second purpose is the execution of the power stroke[20]. The LN is capable of bending at its center which serves as a hinge. It can therefore be found in one of two conformations - extended, and bent. The transition between the bent conformation and the extended conformation is considered as the power stroke and is controlled by the nucleotide state of the AAA1 unit (Fig. 2.2)[20,46]. In addition to these two main functions, the LN appears to play an important role in the regulation of the chemomechanical cycle of the motor and certain studies indicate that it is involved in a gating mechanism of the motor[47]. In particular, these studies show that the application of an external load to the LN modulates the affinity of the motor for MT, and therefore the LN is important in the context of communication between the two MDs in the dimer.

One of the surprising and interesting facts about dynein is that its MTBD is not located near the ATP binding site but rather at the end of the stalk, a longer than 10nm coiled coil structure, placing it more than 20 nm away from the AAA1 unit[18]. It has been shown that not only does the AAA1 unit control the MTBD's affinity for MT, the presence of MT can affect the nucleotide affinity and hydrolysis rate at the AAA1 site[18,43,48,49,50]. Therefore, the natural question to ask is: what is the allosteric mechanism behind the communication between two domains that are

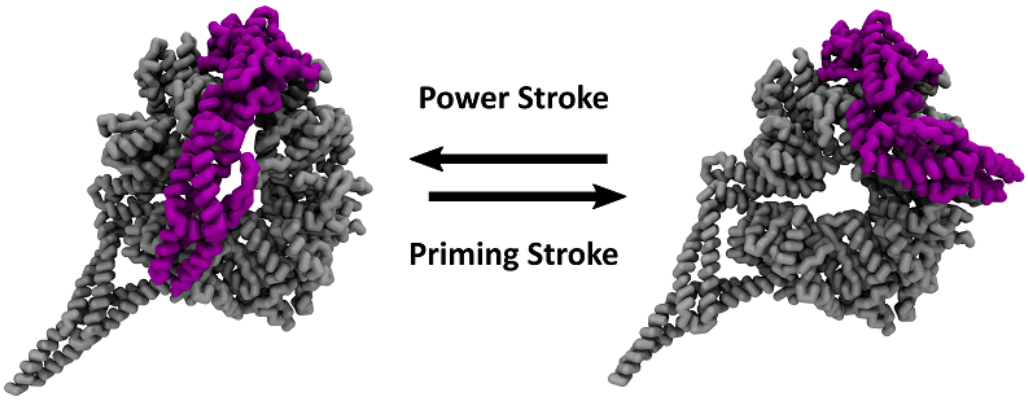


Figure 2.2: Conformational transitions of the LN domain. The transition from the bent conformation (right) to the extended conformation (left) is referred to as the power stroke. The reverse transition is referred to as the priming stroke.

so far apart? The current consensus is that the information to and from the MTBD is transferred through the stalk by a relative shift in the positions of the coils in the coiled coil structure. The domain that seems to play an important role in this process is the strut, a helical structure that connects the AAA5 unit with the stalk and appears to be involved in the allosteric communication[51,52,53].

Thanks to several studies that measured the kinetics of the enzymatic processes in the dynein motor we now have a much better understanding of its chemomechanical cycle[41,49,50]. Setting the AAA3 and AAA4 sites aside, the cycle is predominantly controlled by the AAA1 sub domain. When the AAA1 catalytic site is empty, also known as the apo state, the motor binds tubulin with high affinity and does not detach. The LN domain is fully extended at this state with the N-terminal attached to the AAA5 unit[17]. Binding of ATP at the AAA1 site induces two events. The first is a rapid dissociation from the MT[41]. The second is a fast conformational transition of the LN from the extended or post-stroke state to the bent or pre-stroke

conformation[46,49]. This transition is referred to as the priming stroke and is a precursor of the power stroke (Fig. 2.2). Eventually the ATP in the AAA1 site is hydrolyzed into ADP and a phosphate (Pi). It has been shown that upon the Pi release, the affinity of the dynein motor for MT increases and the motor rebinds to tubulin rapidly[41]. In addition to binding MT, the Pi release is thought to initiate two other processes. The first is the release of the ADP molecule bound to the AAA1 site. The second is the power stroke, during which the LN transitions from the pre (bent) to post (straight) stroke conformation[46,54,55]. Both these processes are quite slow. However, it was demonstrated in several studies that the presence of MT dramatically accelerates these processes, just as in kinesin[48,49,56]. As a result, it is unlikely for the power stroke to occur before the motor rebinds the MT. Upon the ADP release the cycle is complete and can be repeated.

The last decade has witnessed significant advancements in the measurement of the "macroscopic" properties of the dynein motor. Advances in single molecule experimental techniques have enabled us to track individual dynein motors. Like its counterpart kinesin, dynein is a processive motor with run lengths on the order of magnitude of micrometers[14,39]. Though it is slower than kinesin, it can reach velocities as high as 200 nm per second[13,14,39]. The most interesting measurements, however, are those of the step size of the motor. While kinesin takes steps of 8 nm in length with minor variations, dynein has a wide step size distribution with an irregular shape[13,14,39]. Furthermore, it is not uncommon for the dynein motor to step backwards towards the positive end of the MT even when no external resisting forces are exerted on the motor. Yet another important difference between the two motors is that while kinesin tends to walk along a single protofilament of the MT, dynein readily and regularly takes side steps and hops from one protofilament to the next[13,14]. Understanding the stepping pattern of dynein theoretically is therefore a difficult problem as it involves at least two dimensions, even if stepping is largely

controlled by diffusion.

The two dimensionality of the stepping pattern raises questions concerning the communication between the two dynein motors. Several studies show that the internal strain between the two MDs regulates the stepping behavior of the motor[13,47,57,58]. More specifically, the direction with which the strain is applied determines which MD will step next and which MD will remain bound to the filament. This is often referred to in the literature as gating. In the case of kinesins and myosins, it is easy to define which MD is leading ahead and which MD is trailing behind, due to the fact that these motors walk along a single filament. The direction of the force acting on each MD is therefore easy to determine. In dynein, however, since each MD can bind a different protofilament, it is possible in principle to find both MDs adjacent to each other. It is impossible to determine which MD is leading and which MD is trailing in such a scenario. Therefore, it is unclear how the two MDs coordinate their movement. There are several experimental results that hint at the answer to this question. As the distance between the two MDs in their bound state increases, the likelihood of the leading MD to step increases[13,14]. In other words, the further the MDs are apart, the more coordinated the stepping pattern becomes. Furthermore, recent studies indicate that the LN plays an important role in the coordination and gating mechanism of the two MDs. In particular, it has been demonstrated that a pulling force applied to the N-terminal of the LN results in a lower detachment rate when compared with force applied to the C-terminal domain[47]. This serves as strong evidence of the involvement of the LN in the regulation of the activity of the motor.

It is clear that dynein is a sophisticated machine with a complex mechanism of action. Whether it is how the dynein stepping pattern emerges from the underlying architecture of the motor, or what are the allosteric pathways that are involved in the motors domain chemomechanical cycle, many questions remain open with

respect to its function. Our goal in this work is to address some of these questions as quantitatively as possible by devising theories and computations that reproduce many of the features observed in experiment.

# Chapter 3

# Theory for the Stepping Mechanism of Dynein

## 3.1 Introduction

Due to the high level of complexity of the dynein motor, it is a major challenge to make a connection between the molecular mechanisms and the measurable properties of the motor. At the level of a single MD, there is a sophisticated interplay between ATP hydrolysis and several processes including conformational transitions and changes in the affinity of the motor for MT[41,43,45,49]. While the underlying mechanisms of these processes are hard to explain, their outcomes can be described more simply.

Over the last several decades there has been a growing number of attempts to tackle the motor protein problem using different theoretical approaches. One such approach is to model the chemomechanical cycle of the motor using a simple jump process between discrete states and coupling the cycle to the motor's step[59,60,61,62]. This type of model has been extremely successful at reproducing experimentally measured values of properties such as dwell times, velocities and run lengths. Fur-

thermore, these models can predict the response of the motor to an external force or to changes in ATP concentration. Kinesins in particular, lend themselves to this type of models due to the fact that their step size is very well defined and shows little to no variability[63].

This approach has proven to be at least partially successful for larger motors such as dynein, however, these systems are more complex due to the fact that the step size distribution has a large variance[39]. While taking into account only forward steps under zero load and both backward and forward steps under a stalling force seems to be sufficient for kinesin, that is not the case for dynein.

High speed atomic force spectroscopy studies have shown that Myosin V (MV) can "stomp" in place, that is, one of the MDs detaches and then rebinds to the same site on the actin filament. This stomping motion is not expected to affect the step size distribution since zero sized steps are not experimentally detectable. On the other hand, it is expected to influence the dwell time between steps and therefore has to be taken into account when modeling the chemomechanical cycle.

For dynein the problem is further complicated by the fact that the step size distribution does not have a simple shape. In addition, dynein regularly performs backwards steps even under zero load[13,14,39]. Another feature that distinguishes dynein from kinesin is that while these two motor super-families walk along a single filament, dynein can hop from one protofilament to the next[13,14,64]. If that is not enough, the two MDs in dynein step in a partially uncoordinated fashion[13]. In other words, at each step, either of the two MDs is likely to take the role of the stepping domain.

Due to the complex nature of these systems, a simple model of the step mechanics is not enough to account for the motor's properties and a more elaborate model of both the chemomechanical cycle and of the stepping mechanism itself is required. There have been a few attempts to tackle the particular problem of the step

size distribution of dynein using elaborate models[65,66]. These models do succeed in reproducing certain features of the step size distribution but with a few caveats. First, the step size distributions obtained in these studies were discrete. This is in contrast to the experimentally obtained distributions which seem to behave in a continuous manner. This is due to noise introduced by thermal fluctuations and measurement error. At the molecular level we expect dynein to step between discrete binding sites. However, in order to properly compare a model to experiment, noise has to be taken into account. Second, these studies do not successfully couple the mechanical step to a chemomechanical cycle unless some of the rate constants deviate significantly from their experimentally determined values. Third, to our knowledge, these studies resorted to numerical calculation in order to obtain their step size distributions. Since numerical calculations, by their nature, are slow, it is hard to properly explore the parameter space without lengthy computation times.

Here, we present a novel theoretical approach using an analytically solvable model for the mechanical step of dynein, combined with a kinetic model for the chemomechanical cycle. Our model, based on a previous work on Myosin V, is able to successfully reproduce a variety of properties of the dynein motor, including step size and dwell time distribution. The extracted parameters have a clear physical interpretation with values that correspond to the experimentally determined range[67]. We are also able to make predictions as to how the behavior of dynein would change under the variation of certain conditions, such as ATP concentration.

## 3.2 Mechanical Model

The first part of our model is a simplified representation of the MD dimer that accounts for the stepping mechanism. We model a single MD as a rigid rod, pointing from the MTBD to the N-terminus of the LN. The rod is characterized by the end-

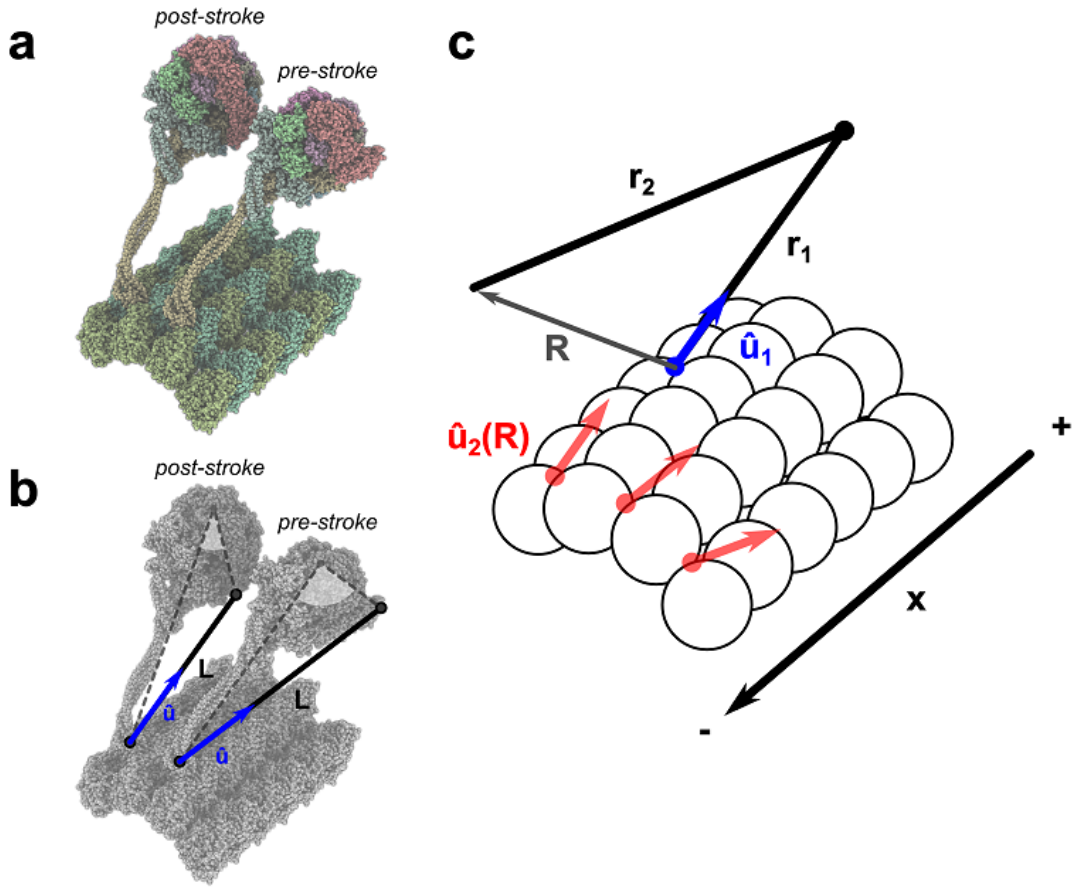


Figure 3.1: Mechanical model of the dynein motor. a) Two MDs bound to the MT, one in the post-stroke conformation and the other in the pre-stroke conformation. b) The rigid rod is defined by stretching a line from the motor's MTBD to the N-terminal of the LN. The blue vectors represent the directional constraint $\hat{\mathbf{u}}$ acting on the rod. The length of the rod, L, as well as the direction of the constraint depend on the conformational state of the MD. c) The dimer composed of two rods, represented by the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, joined by a freely rotating joint. $\mathbf{R}$ is the vector joining the MTBD ends of the rods. The red vectors represent the directional constraints that affect the binding affinity of the stepping rod towards the different TBSs. As can be seen, each TBS has a different orientation preference, making $\hat{\mathbf{u}}_2$ a function of $\mathbf{R}$.

to-end vector, $\mathbf{r}$, with an associated length, $|r| = L$. When bound to the MT, the stalk domain is forced to adopt a particular orientation. We account for this effect by imposing a constraint on the rigid rod such that $\mathbf{r}$ is more likely to point along a constraint vector $\hat{\mathbf{u}}$ (Fig. 3.1b). The strength of the constraint is determined by an effective parameter, $\mathcal{T}$. The equilibrium probability distribution for $\mathbf{r}$ when the MD is bound to the MT is given by,

$$\mathcal{P}(\mathbf{r}) = \frac{\mathcal{T}}{\sinh \mathcal{T}} \frac{\delta(r - L)}{4\pi L^2} \exp(\mathcal{T} \hat{\mathbf{u}} \cdot \hat{\mathbf{r}}) \tag{3.1}$$

To address the stepping mechanism, we assume that each MD is in one of two conformations, the extended pre-stroke and the compact post-stroke states (Fig. 3.1a). The transitions between these two states, when coupled to the chemical cycle of the motor, allow dynein to propel itself forward along the MT. There are two main structural differences between the two mechanical states (Fig. 3.1b). The first is the preferred orientation with which the stalk domain aligns itself with respect to the MT. The second is a large change in the orientation of the LN with respect to the rest of the MD. This second structural change is associated with the motor's power stroke and is expressed as a change in L in our model. EM and crystallographic studies show that within each of the two states the LN maintains an approximately constant position, thus justifying our use of a rigid rod model with constant length [19]. We obtained the values for L and the orientation vector $\hat{\mathbf{u}}$ for each state by fitting the corresponding dynein crystal structures to that of a MTBD bound to the MT (see appendix A) [18,19,21].

With the goal of reproducing the experimentally measured step size of the motor in mind, we are interested in modeling the behavior of a dimer of MDs. We know that the two MDs dimerize through the N-terminals of their LNs. Therefore, we treat the dimer as a couple of rigid rods connected at their LN ends by a freely rotating joint (Fig. 3.1c). Typically, a step starts with both MDs bound to the MT.

Then, one of the MDs detaches from the MT. We designate this MD as the Stepping Motor (SM). The SM diffuses, while tethered to the bound MD, which is the Non-Stepping Motor (NM). Finally, the SM rebinds to the MT at a different binding site further towards the (-) end of the MT. How likely is it for the SM to bind to each of the geometrically available binding sites along the MT? By answering this question we can calculate the step-size distribution.

In order to answer the question posed above we make an additional approximation. We assume that the stepping time, the time that passes from the moment the SM detaches from the MT until it rebinds at the Target Binding Site (TBS), is significantly longer than the microscopic relaxation time of the rotational diffusion of the SM's orientation. In other words, we assume the orientation of each MD reaches equilibrium before the step is completed. Such an assumption is reasonable since attachment to the MT requires ATP hydrolysis and phosphate release, which occurs in a time scale of milliseconds[18,40,43,46]. With this assumption, the probability of the SM binding to a particular TBS is given by the probability of finding the free MTBD in the neighborhood of the TBS in equilibrium.

In terms of our model, if $\mathbf{r}_1$ represents the NM and $\mathbf{r}_2$ represents the SM, the problem can be reformulated to finding the equilibrium probability distribution of the vector $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2$, which is the vector pointing from the NM's MTBD to the SM's MTBD (Fig. 3.1c). This distribution can be obtained using,

$$\mathcal{P}(\mathbf{R}) = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \mathcal{P}(\mathbf{r}_1)\,\mathcal{P}(\mathbf{r}_2)\,\delta(\mathbf{R} - \mathbf{r}_1 + \mathbf{r}_2) \tag{3.2}$$

with $\mathcal{P}(\mathbf{r})$ given by Eq. 3.1. It is important to note at this point that the angular constraints in $\mathcal{P}(\mathbf{r}_1)$ and $\mathcal{P}(\mathbf{r}_2)$ have different physical interpretations. In the case of $\mathcal{P}(\mathbf{r}_1)$ the constraint represents the actual angular tension imposed on the NM by the MT. In the case of $\mathcal{P}(\mathbf{r}_2)$ the interpretation is more subtle. Since the SM is not bound to the MT, there is no angular force acting on the stalk. That said,

the SM cannot bind to the TBS at any orientation and is limited to those that optimize binding between the SM and the MT. Therefore, our constraint represents the tendency of the SM to bind to the TBS at the optimal orientation. Consequently, the constraint vector $\hat{\mathbf{u}}_2$, acting on $\mathbf{r}_2$, is a property of the individual TBS and therefore a function of $\mathbf{R}$.

We can solve Eq. 3.2 analytically to obtain,

$$\begin{aligned} \mathcal{P}(\mathbf{R}) = 2\pi N L_1^2 \sin\theta_1 \exp\left(-\mathcal{T}\hat{\mathbf{u}}_2 \cdot \frac{\mathbf{R}}{L_2} + \mathcal{T}_c \cos\theta_c \cos\theta_1\right) \\ \times \\ I_0\left(\mathcal{T}_c \sin\theta_c \sin\theta_1\right) \end{aligned} \tag{3.3}$$

where $I_0$ is the modified Bessel function of the first kind of order zero, $\theta_1$ is the angle between $\mathbf{R}$ and $\mathbf{r}_1$, and $\theta_c$ is the angle between a linear combination of the constraint vectors and $\mathbf{R}$.

Our expression for $\mathcal{P}(\mathbf{R})$ allows us to calculate the probability of the SM reaching a particular binding site on the MT. Our goal, however, is to calculate the step size probability distribution, requiring knowledge of the binding site to which the SM is bound initially. In particular, we are interested in the projection of the displacement vector along the $\hat{\mathbf{x}}$ (MT) axis.

In order to calculate the motor step size we need to consider two scenarios. In the first scenario the same MD steps twice consecutively. If we define the position of the initial binding site of the SM as $x_i$ and the final position as $x_f$, then the step size is given by,

$$\Delta x_{i,f} = x_f - x_i \tag{3.4}$$

As mentioned above, once stepping starts the MD equilibrates rapidly and looses memory of the initial conformation. Thus, the position of the final binding site is independent of the position of the initial binding site. Therefore, the probability

distribution of the step size along the $\hat{\mathbf{x}}$ axis is,

$$\mathcal{P}(\Delta x) = \sum_{n}^{N} \sum_{m}^{N} \mathcal{P}_f(\mathbf{R}_n)\, \mathcal{P}_i(\mathbf{R}_m)\, \delta(\Delta x - (x_n - x_m)) \tag{3.5}$$

where N is the number of potential binding sites accessible to the dynein motor and $\mathcal{P}_i$ and $\mathcal{P}_f$ are the probabilities for the initial and final binding sites respectively.

In the second scenario the two motors step one after the other in an alternating manner. In this case, we define the position of the motor that stepped last as the new origin and $x_i$ becomes $-x_i$. The step size is given by

$$\Delta x_{i,f} = x_f + x_i, \tag{3.6}$$

and the step size distribution is,

$$\mathcal{P}(\Delta x) = \sum_{n}^{N} \sum_{m}^{N} \mathcal{P}_f(\mathbf{R}_n)\, \mathcal{P}_i(\mathbf{R}_m)\, \delta(\Delta x - (x_n + x_m)) \tag{3.7}$$

This distribution is discrete in space since there is a discrete number of potential binding sites at specific locations. The experimentally measured step size distribution, however, seems to be continuous because the detection of the quantum dot position introduces errors. In order to obtain a distribution that is comparable with experiment we add Gaussian "noise" to our distribution.

## 3.3 Kinetic Model

The probability distribution $\mathcal{P}(\mathbf{R})$ in Eq. 3.3 depends on the conformations of both MDs. Furthermore, the step size probability distribution depends on the likelihood that the same MD steps twice, or alternatively, that the two MDs step in an alternating fashion. In order to calculate these probabilities we need a kinetic model

that couples the structural transitions to the different states in the chemical cycle of each individual MD as well as the dimer as a whole.

### 3.3.1 Chemomechanical Cycle

We first define the chemomechanical cycle of a single MD (Fig. 3.2). In the initial state, which we designate as state 1 in Fig. 3.2, the MD is tightly bound to the MT, with ADP bound to the AAA1 and the LN in the extended pre-stroke conformation.

**Power Stroke.** Experiments show that when the MD is in the extended conformation (pre-stroke) and with ADP in the AAA1 binding pocket, the LN undergoes a power stroke in which its conformation changes from extended to a compact state[41,46]. Eventually, ADP is released. Furthermore, it has been shown that binding of the motor to the MT accelerates both these processes[49]. The findings lead us to postulate that the power stroke and ADP release are coupled, permitting us to treat them as a single process. We designate the rate of this step as $k_p$. The MD then transitions to the post-stroke conformation Apo state, designated as state 2 in Fig 3.2.

**ATP binding.** Once ADP is released from the AAA1 module, ATP can bind at a rate $k_{ATP}$. We assume that this binding rate is linear in ATP concentration in the concentration range of interest. We designate the MT bound post-stroke conformation with ATP in the AAA1 site as state 3 (Fig 3.2).

**MT dissociation.** ATP binding to the AAA1 unit has been shown to induce rapid dissociation of the dynein motor from the MT[41]. The dissociation rate is $k_d$. However, recent studies indicate that this process is highly dependent on the nucleotide state of the AAA3 module[43,44,45]. When ADP is bound to the AAA3 subunit with ATP bound to the AAA1, rapid MT dissociation does occur. However, when the AAA3 unit is either occupied by ATP or empty, the dynein motor remains bound to the MT regardless of the nucleotide state of the AAA1 unit. The exception

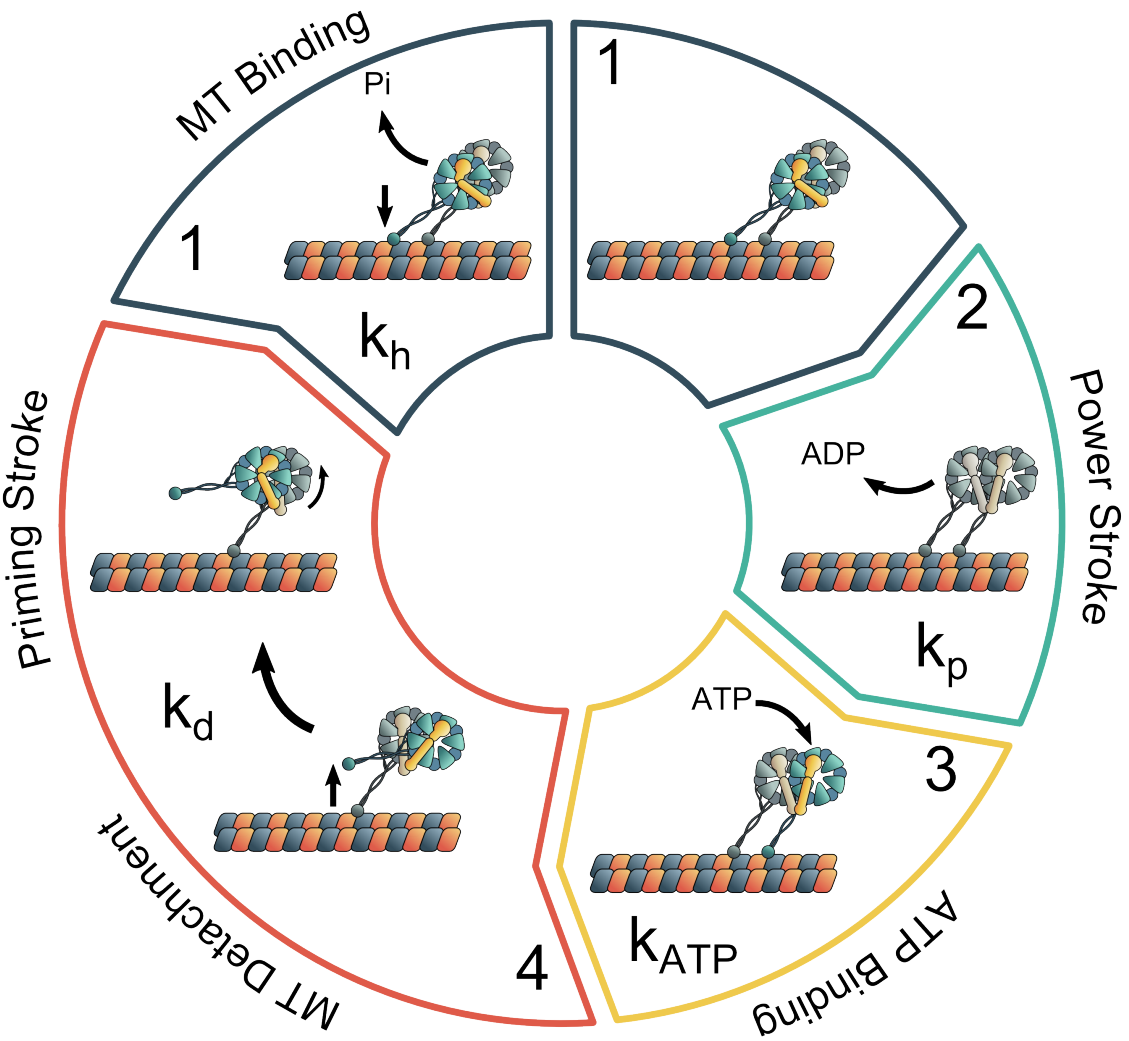


Figure 3.2: The chemomechanical cycle of a single MD. State 1: the motor is in the pre-stroke conformation, bound to the MT with ADP in its AAA1 module. Transition 1-2: the MD undergoes a power-stroke and releases ADP. Transition 2-3: The MD binds ATP. Transition 3-4: the MD detaches from the MT and simultaneously undergoes a priming-stroke. Transition 4-1: The MD hydrolyzes ATP, releases the phosphate and binds to the MT, completing the cycle.

to this rule is when tension is applied to the LN. It has been shown that even when there is ATP in the AAA3 unit, ATP binding to the AAA1 unit induces detachment from the MT as long as tension is applied to the LN domain[45]. Once the dynein motor detaches from the MT, ATP at the AAA1 site induces a priming stroke of the LN[41]. The priming stroke, returning the LN to the extended state, occurs rapidly. We therefore assume that a SM is always found in the pre-stroke conformation. The unbound pre-stroke conformation with ATP in the AAA1 site is state 4 (Fig 3.2).

**Phosphate release and MT binding.** The detached MD does not regain a high affinity for MT until the ATP at the AAA1 site is hydrolyzed and the inorganic phosphate (Pi) is released[41]. Once the Pi detaches, dynein binds rapidly to the MT. In our model, we group these three processes, hydrolysis, Pi release, and MT binding as a single step. We make the assumption that the MT binding rate is significantly faster than the hydrolysis rate, and therefore we can neglect it in our calculations. The rate of hydrolysis/phosphate release has been measured experimentally, $k_h$[18,40,43,46]. Once dynein binds to the MT the system returns to state 1 and the cycle is complete.

### 3.3.2 Stepping States and Coordination

The distribution $\mathcal{P}(\mathbf{R})$ is determined by the mechanochemical states of both MDs in the dimer. Since in our model the SM is always in state 4 and the NM can be found in states 1,2, and 3, there are only three possible stepping states. We designate the corresponding states as A,B, and C accordingly. We assign letters to the stepping states of the MDs in the dimer to distinguish them from the mechanochemical states of the individual MDs. In state A, both the NM and SM are in the extended pre-stroke conformation and, therefore, forward and backward steps occur with almost equal probability. Thus, mode A is a non advancing mode. Fig. 3.3a shows the corresponding $\mathcal{P}(\mathbf{R})$ at the different TBSs along the MT. In states B and C the

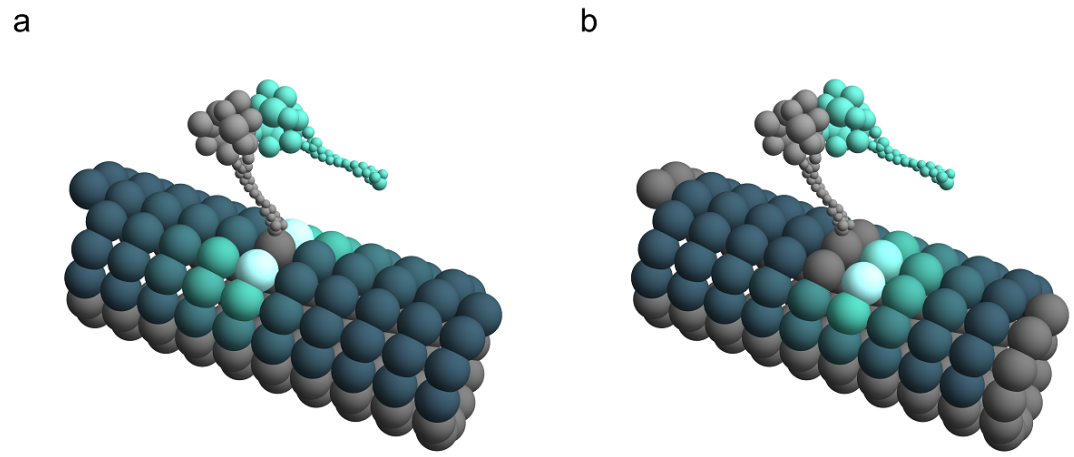


Figure 3.3: The probabilities of binding to different TBSs on the MT. The lighter the color of a particular TBS, the more likely the SM is to bind to that TBS during a step. a) The TBSs binding likelihoods during stepping state A. b) The TBSs binding likelihoods during stepping states B/C. The SM tends to bind on either side of the NM in state A while in states B/C the SM tends to bind in front of the NM.

bound motor is in the contracted post-stroke form while the SM is in the extended conformation resulting in a preference for steps towards the minus end of the MT. Thus, states B and C are advancing states with identical $\mathcal{P}(\mathbf{R})$. The difference in binding pattern between the different states can be seen by comparing Fig. 3.3a and 3.3b. According to the stepping pattern in state A the SM tends to bind on either side of the NM without a significant preference for the front or back. The stepping pattern of states B and C, in contrast, shows a clear preference for binding sites in front of the NM. It is important to note that the binding pattern of states B and C is not entirely symmetrical with respect to the MT axis when looking from above. In the case in which both motors are found in state 4 the dynein dimer simply falls off the MT and the run is terminated. We designate this termination by $T_{er}$.

Several single molecule studies indicate that the two motors partially coordinate their steps[13,14]. Furthermore, previous theoretical studies argued that inter-motor coordination is necessary for dynein to simultaneously maintain both high

processivity and high velocities[62]. It has been suggested that the AAA3 module is involved in the regulation of dynein activity but interestingly, its control of MT affinity in the presence/absence of tension on the LN makes it a perfect candidate for a gating mechanism[43,45].

To simplify the theory we assume that the AAA3 module can be in one of two states: open and closed. When the AAA3 module is in the open state, binding of ATP to the AAA1 site will lead to rapid disassociation from the MT. In the closed state the MD remains bound to the MT even when ATP binds to the AAA1 module. The exception to this is when tension is applied to the LN. In that case the AAA3 module seems to retain an open form regardless of its nucleotide state. Since there is no tension acting on the LN when the SM is unbound, the NM cannot detach from the MT if its AAA3 site is closed and therefore the run cannot be terminated. In other words, the higher the likelihood of the AAA3 module being in the closed state, the lower the probability that the run will terminate and the higher the processivity is. We assume that the AAA3 ATPase cycle is independent of the AAA1 cycle and that steady state has been reached[44]. Under these conditions the probability of the AAA3 site of either MD being open is constant in time and we designate it by $\gamma$.

Since we expect $\gamma$ to be small, given the high processivity of dynein, there has to be a tension generating mechanism to ensure that the motor maintains motility. As there is no tension acting on the LN during the stepping process itself, we expect that immediately after the SM binds to the MT no tension is added. This implies that the tension is generated by the power stroke of either motor. To simplify the analysis we make the following approximation: there is tension acting on the LNs when both MDs are bound to the MT only if one or both MDs underwent a power stroke.

### 3.3.3 Markov Chain Model

We are interested in calculating the probability with which each stepping state occurs as well as the dwell times between the different steps. Additionally, we want to calculate the termination probability. In order to calculate these probabilities, we construct a Markov chain for the transitions between the different stepping states. To do so, we determine the mechanochemical state of each MD at each stepping state and then construct a kinetic scheme to determine all possible transition pathways that lead to the next stepping state (Fig. 3.4). Given the rates of the transitions between the mechanochemical states we can also obtain the dwell time probability distribution for each transition.

Given that there are three possible stepping states for each MD and either MD could detach to be the SM, we need to consider six different stepping states in our Markov model. If we designate the MDs as 1 and 2 then we can define the set of possible stepping states as $S_n = \{A_1, B_1, C_1, A_2, B_2, C_2\}$ where $n$ is the step number over time and the subscript indicates which of the two motors performed the step. Our Markov transition matrix can now be obtained by calculating the conditional transition probabilities between the different stepping states:

$$M_{i,j} = \mathcal{P}\left(S_{n+1}\left(j\right)|S_n\left(i\right)\right) \tag{3.8}$$

Since dynein is a processive motor, the probability of run termination is small. Under these conditions it is not unreasonable to assume that the system reaches a steady state and that the probability of the motor stepping in a particular stepping state is time independent. Eq. 3.8 can be then written as,

$$M_{i,j} = \mathcal{P}\left(S_j|S_i\right) \tag{3.9}$$

The probability $\mathcal{P}\left(S_i\right)$ can be found by finding the eigenvector of M with

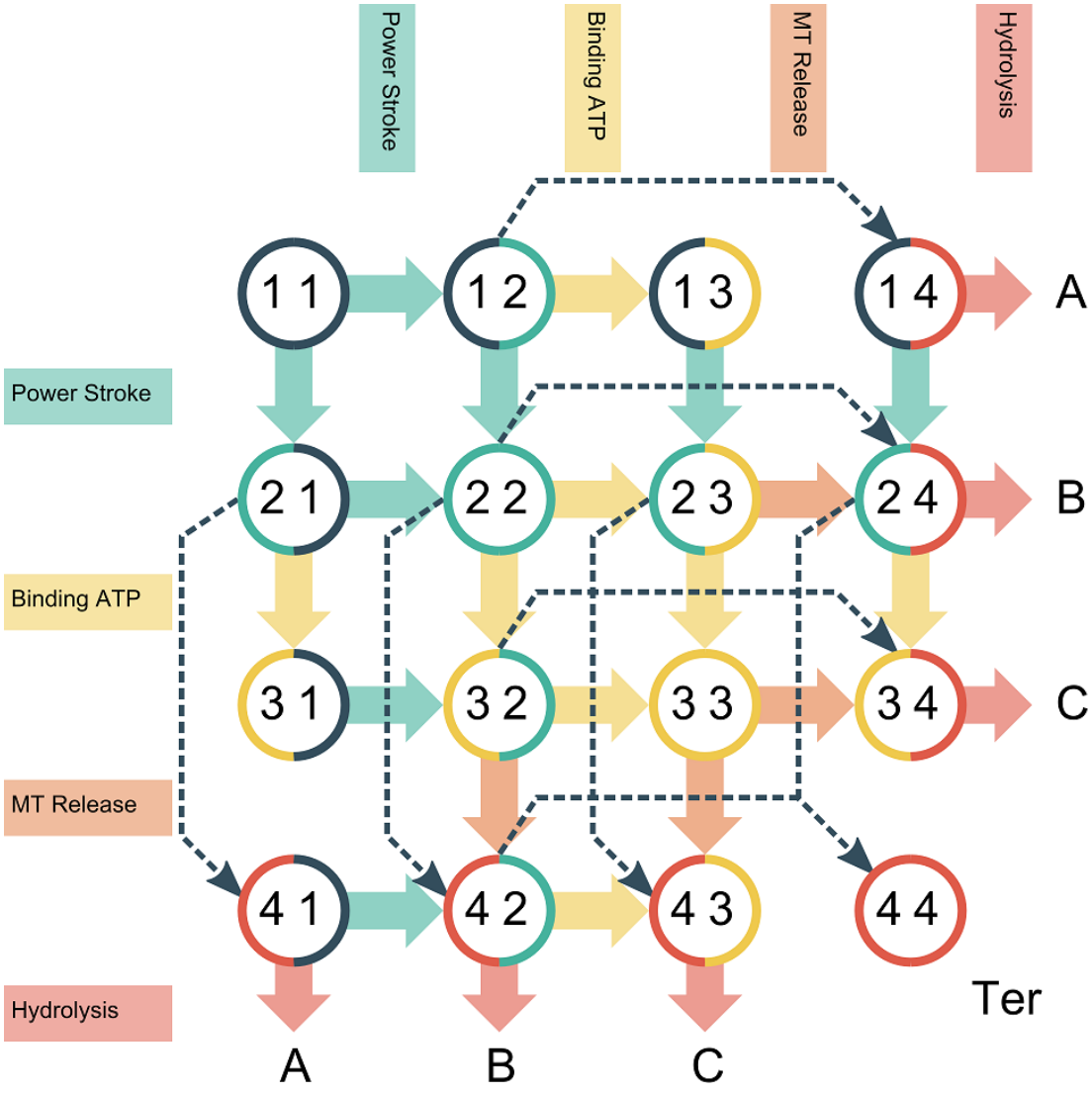


Figure 3.4: The kinetic scheme for the MD dimer. The two numbers in each circle are the states in the mechanochemical cycle of each of the two MDs. Moving downwards in the chard represents the progression in the mechanochemical cycle of MD number one. Moving towards the right in the chard represents progression in the cycle of MD number two. Depending on the transition rates, the probabilities of the dimer stepping in states A, B, and C or falling of the MT can be calculated. The dashed lines represent transitions that are possible only when the AAA3 module is in the open state. The probability of following such a transition is given by $\gamma$. The initial position in the scheme is determined by the row/column in which the previous step ended.

eigenvalue unity. The probability of termination can be found if we consider dynein detaching completely from the MT as a seventh absorbing state in our model. Assuming the run termination is a Poisson process, we can calculate the mean number of steps taken by the motor by taking the reciprocal of the termination probability. By multiplying by the mean step size we obtain the mean run length, which can be compared with experimental measurements.

In addition to finding $\mathcal{P}(S_i)$, we are interested in the mean dwell time between steps as well as the dwell time distributions. We therefore construct two additional matrices. The first matrix is $T_{i,j} = \tau(i,j)$ whose elements are the mean dwell times for each transition in M. The second is $F_{i,j} = f_{i,j}(t)$ whose elements are the dwell time distribution functions for each transition in M.

Typically, in single molecule studies of dynein, the motor is labeled using a quantum dot. There are two common variations on this procedure. The first is when one of the two MDs in the dimer is labeled (MD labeling). The second is when the tail domain between the two MDs is labeled (tail labeling). The nature of the step size and dwell time distributions obtained in such studies depends on which type of labeling was used. Since our mechanical model tracks the position of the MDs, we compare our results to MD labeled measurements for both step sizes and dwell times.

The dwell time between steps of a single MD as well as the step size depend on the probability of the same MD stepping more than once consecutively, or alternatively on the probability that the MDs alternate their steps. We designate the conditional probability that the last MD to step will be the next SM as $P_c$ (consecutive steps) and the probability that MDs will alternate in their steps $P_a$ (alternating steps). In a similar manner we define $\tau_c$ and $\tau_a$ as the mean dwell times between consecutive and alternating steps respectively and $f_c(t)$ and $f_a(t)$ as the corresponding dwell time distributions.

By considering all possible combinations of consecutive and alternating step, we obtain the following expressions for the MD mean dwell time and dwell time distribution:

$$\tau_{MD} = P_c\tau_c + \frac{2\tau_a + P_c\tau_c - 2\tau_a P_c}{(P_c - 1)^2} \tag{3.10}$$

$$f_{MD}(t) = P_c\tau_c + \mathcal{L}^{-1}\left(\frac{P_a^2 \tilde{f}_a(s)^2}{1 - P_c\tilde{f}_c(s)}\right) \tag{3.11}$$

where $\tilde{f}_c(s)$ and $\tilde{f}_a(s)$ are the Laplace transforms of $f_c(t)$ and $f_a(t)$ respectively and $\mathcal{L}^{-1}$ is the inverse Laplace transform.

## 3.4 Results

### 3.4.1 Comparison with experiment

We fit the four parameters in our model in order to reproduce the motor's velocity and processivity at saturating ATP concentrations as well as the mean step size and mean dwell time in low ATP concentrations[13,14]. The values of the parameters are listed in Table 3.1.

Fig. 3.5a shows a comparison between the experimental and the calculated step size distribution of a MD labeled dynein. Remarkably, even though we used only the mean step size in our fit, our model nearly quantitatively reproduces to a good degree the shape and features of the experimental distribution.

In order to better understand how this distribution comes about we calculated the contributions to the step size distribution due to both alternating and non-alternating steps. As can be seen from Fig. 3.5b non-alternating steps have a nearly symmetrical distribution with major peaks at 8 nm for both the forward and backward steps while alternating steps produce displacement predominantly

| Parameter | Value | Source |
|---|---|---|
| $k_{ATP}$ | $0.148s^{-1} \times Con.$ | Fit to experiment |
| $\mathcal{T}$ | 4.361 | Fit to experiment |
| $\gamma$ | 0.02156 | Fit to experiment |
| $k_p$ | $39.718s^{-1}$ | Fit to experiment |
| $k_h$ | $120s^{-1}$ | [40] |
| $k_d$ | $460s^{-1}$ | [41] |
| $L_{pre}$ | 23.8 nm | [18,19,21] |
| $L_{post}$ | 16.9 nm | [18,19,21] |
| $\hat{\mathbf{u}}_{pre}$ | $(-0.776, 0.327, 0.54)$ | [18,19,21] |
| $\hat{\mathbf{u}}_{post}$ | $(-0.649, 0.016, 0.76)$ | [18,19,21] |

Table 3.1: List of parameters in analytical model.

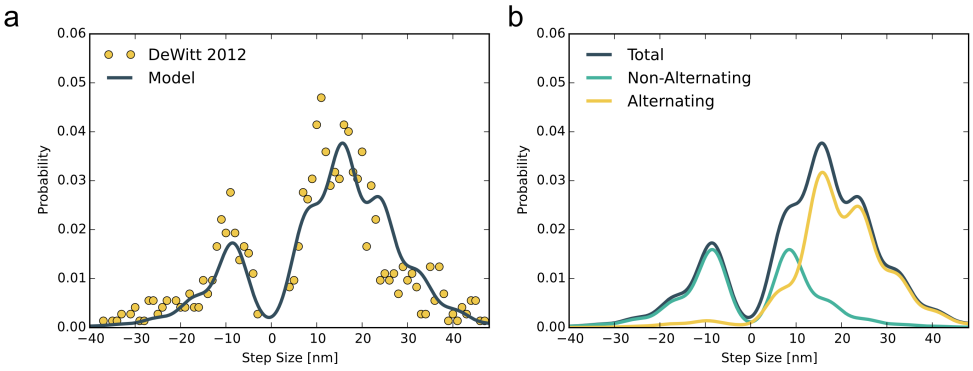


Figure 3.5: The step size distribution of the MD labeled dynein. a) Comparison between the step size distribution that was calculated using our model and the one obtained by DeWitt et.al[13]. b) Decomposition of our step size distribution into the contributions by alternating and non-alternating steps.

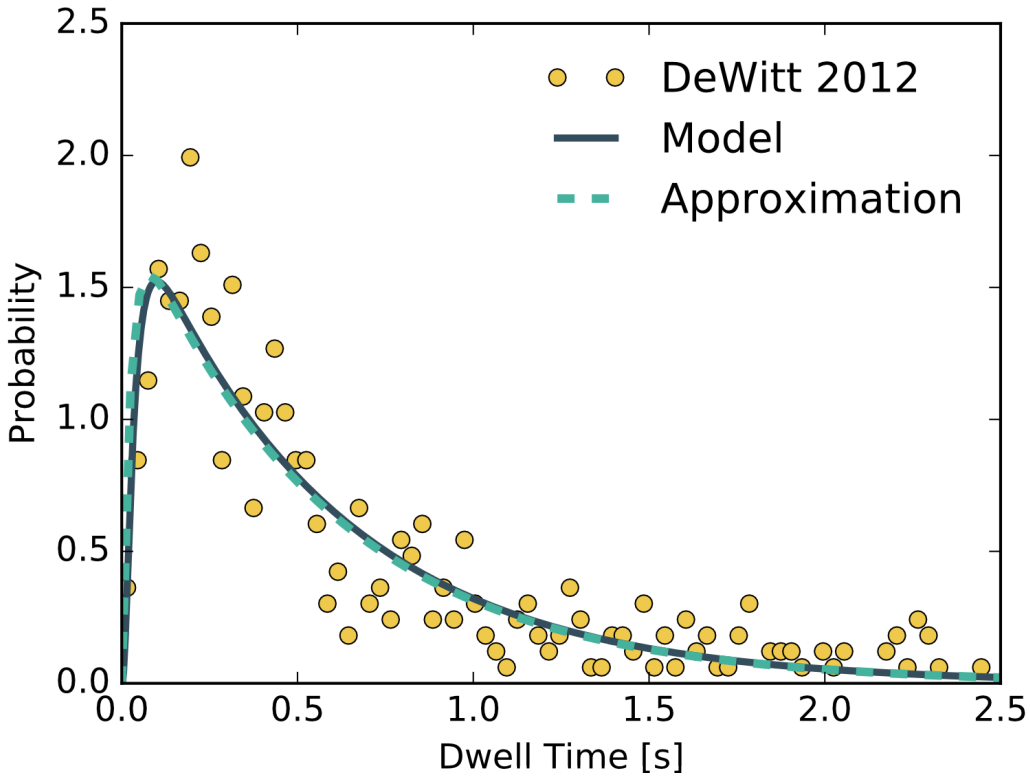


Figure 3.6: The probability distribution of dwell times between steps. The yellow dots represent the experimentally measured distribution while the solid line is the one calculated using our model. The dashed line is the result obtained using our approximation which is valid only under low ATP concentration conditions. As can be seen from the figure, the approximation is almost identical to the result obtained using our model.

in the forward direction. This difference is not surprising. Given that a certain MD steps twice consecutively, the probability distribution of landing in each TBS is identical to the previous step. Therefore, the likelihood of stepping forward or backward with respect to the current TBS is the same. The more interesting finding is that almost all the backward steps are due to non-alternating stepping. In other words, backward steps are due to lack of coordination between the MDs. This is in contrast to kinesin where the steps are fully coordinated and the motor advances in an alternating hand-over-hand motion.

In addition to obtaining the step size distribution we also calculated the dwell

time distribution of a single MD. This can be done using Eq. 3.11 and the result are in Fig. 3.6. Remarkably, our model nearly reproduces quantitatively the experimental distribution well by fitting only to the mean value of the dwell time. The analytical expression for the dwell time distribution obtained from the inverse Laplace transform in Eq. 3.11 is extremely complicated. In order to obtain a simpler expression for the dwell time we note that the experimental dwell time distribution was measured under low ATP concentration conditions and therefore ATP binding is rate limiting. Assuming that the MT detachment and ATP hydrolysis are fast processes, we approximate the mechanochemical cycle as a two state process, powerstroke and ATP binding (Fig. 3.7). We can derive an approximate expression for the dwell time in this cycle:

$$\frac{e^{-k_{ATP}t}\left(1-e^{-k_p t}\right)k_{ATP}\left(k_{ATP}+k_p\right)}{k_p} \tag{3.12}$$

The dashed line in Fig. 3.6 shows that the approximation leading to Eq. 3.12 is excellent, and fits the data well without changing the values of our parameters as reported in Table 1. The main advantage of Eq. 3.12 is that it depends only on two parameters as opposed to the full solution of Eq. 3.11. Another important implication is that the dwell time distribution for a head labeled motor is not simply a convolution of two exponential distributions with identical rates. It is important to note that this approximation relies on the fact that we used a large value for the hydrolysis parameter $k_h$.

Studies have shown that when both the MDs are bound to the MT, the line connecting them does not align along the MT axis and is rather diagonal to it[13,14]. In order to see whether our model can account for this effect we calculated the probability distribution of the angle between the vector connecting the two MDs and the MT axis. This distribution was obtained by calculating the probability of finding each of the MDs at particular TBSs. We then extracted the vectors

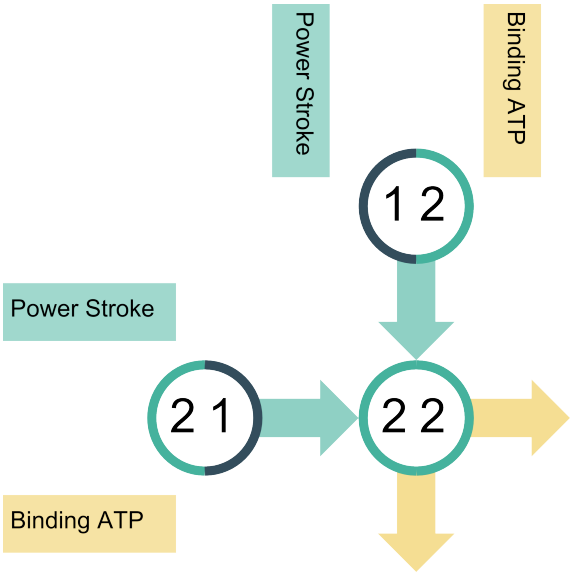


Figure 3.7: An approximation to our kinetic scheme under low ATP concentration conditions. Since the rate of the power stroke and the rate of ATP binding are rate limiting in this scenario, we ignore the rest of the steps in the cycle. For this simplified scheme, reasonable analytical expressions for the dwell time distributions can be obtained.

connecting the two MDs and their corresponding probabilities to obtain an angular distribution. As can be seen from our results (Fig. 3.8), our model does capture the features of the experimentally determined orientation distribution, further validating our theory.

### 3.4.2 Predictions

Our model enables us to make predictions as to how the behavior of dynein will be affected by changing its environment or properties. In particular, we are interested in how changes in the ATP concentration as well as changes in the gating parameter affect the motility of the motor. We first examine the probability of the motor advancing in a particular stepping state under different ATP concentrations and different values of $\gamma$, the probability that the AAA3 site is open. Fig. 3.9 shows that regardless of ATP concentration, the probability of the motor stepping in state A is negligible and the shift occurs only between states B and C. In other words,

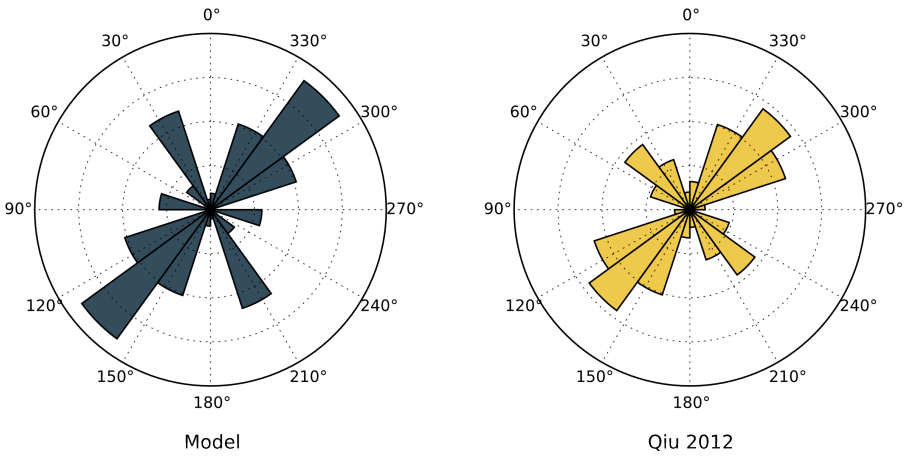


Figure 3.8: The distribution of the orientations between the vector joining the two MDs and the MT axis. The distribution of the left is the one generated by our model and the one on the right is the one obtained by Qiu et.al. As can be seen from the figure, our model reproduces the general features of the experimental distribution relatively well.

even though state A is theoretically possible, it does not occur in practice. This indicates that virtually all backward steps are due to consecutive steps of the same MD and not due to stepping in a non advancing mode. Since states B and C are mechanically identical but chemically distinct, the mean step size in our model is effectively indifferent to changes in the ATP concentration (Fig. 3.10). This implies that the experimentally observed step size distributions, measured in low ATP concentrations, are applicable at high ATP conditions as well. The mean velocity of the motor does increase with ATP concentration because the mean dwell time between steps is shorter. This is simply a consequence of the faster binding of ATP to the MD even though the step size is effectively constant. The overall run length of the motor decreases with an increase in the ATP concentration since the likelihood of each of the MDs to bind ATP and initiate a step during the stepping of its partner increases. Finally, while at low ATP concentrations the probability of the same motor stepping twice consecutively is close to half, this probability decreases with an increase in ATP concentration. This due to the higher probability

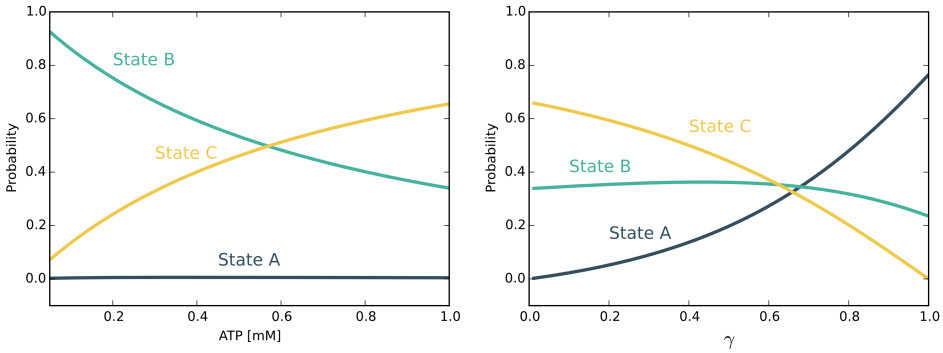


Figure 3.9: The probabilities of the motor stepping in a particular stepping state as a function of ATP concentration and the value of $\gamma$.

of stepping in state C which is more likely to lead to an alternating step.

In contrast to changes in the ATP concentration, increasing the value of $\gamma$ (the probability of the AAA3 site being in the open state) raises the probability of stepping in state A dramatically at the expense of states B and C, as can be seen from Fig. 3.9 Therefore, the more likely it is for one MD to step before its counterpart completes the power stroke. The mean dwell time between steps slightly decreases as we increase $\gamma$ as can be seen from Fig. 3.10. However, the velocity decreases rather than increase because the mean step size dramatically decreases with an increase in $\gamma$. This occurs because at high $\gamma$ values the dominant stepping state is state A which leads to a small step size. Furthermore, the run length becomes small at high $\gamma$ values. A high level of coordination is required for a molecular motor to be processive and therefore, as coordination decreases, so does the motor's run length. As for the probability of the same MD stepping twice consecutively, it decreases with an increase in $\gamma$ since the less coordinated the MDs are, the more likely it is that the same MD will step twice in a row.

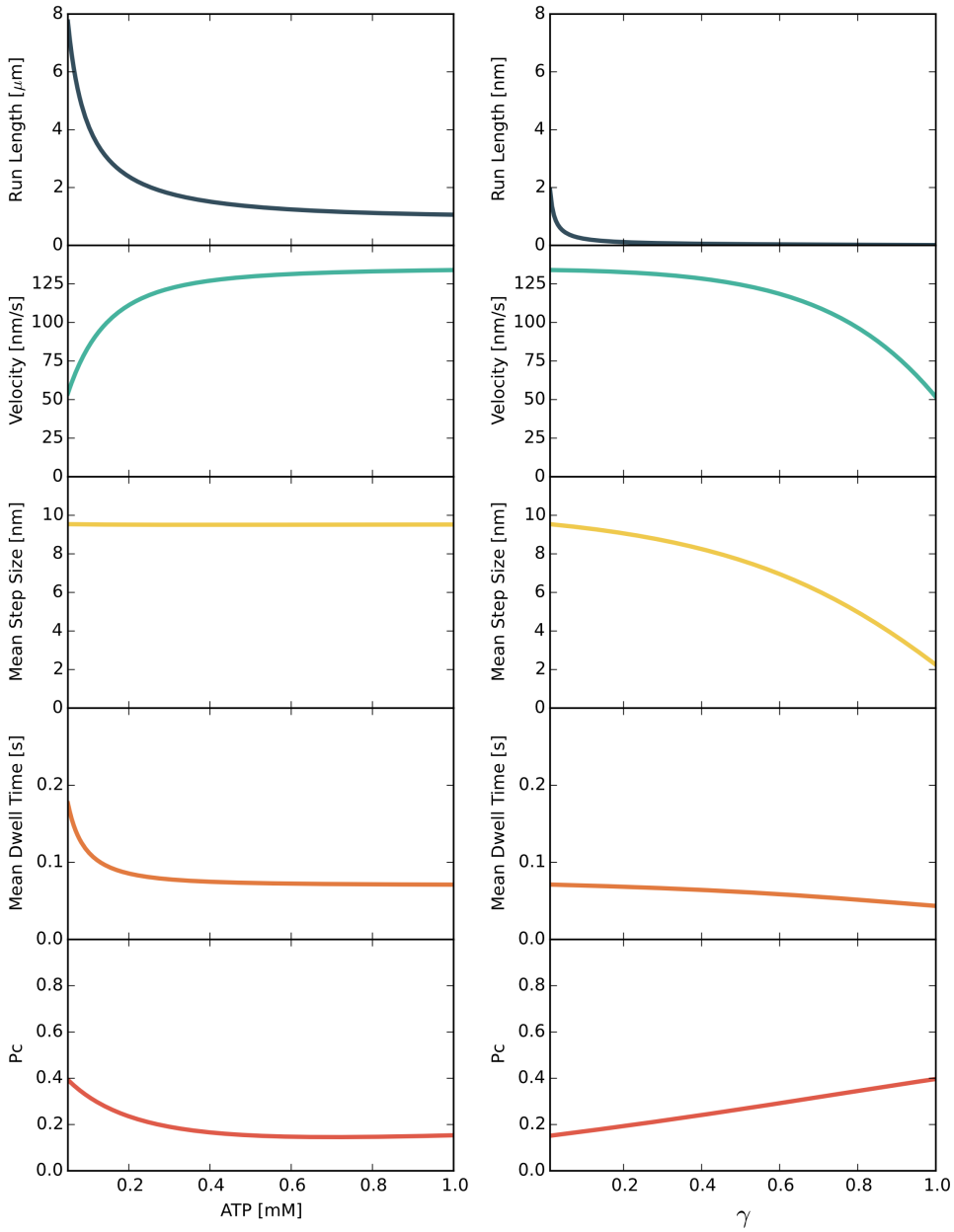


Figure 3.10: Predictions for the values of different physical properties of the motor as a function of ATP concentration and $\gamma$. The observables that were calculated are: Run length, velocity, mean step size, mean dwell time, and probability of consecutive stepping by the same MD (Pc).

## 3.5 Discussion

### 3.5.1 Relationship between ATP and $\gamma$

As mentioned earlier, $\gamma$ can be thought to represent the level of coordination between the MDs or more correctly the lack thereof. How does this coordination arise? As discussed above, the AAA3 module is involved in the coordination between the AAA1 module and the MTBD. Depending on the nucleotide state of AAA3 it allows/inhibits the AAA1-MTBD communication, and therefore is a good candidate for the gating mechanism of dynein. The AAA3 gate is open when it is occupied by ADP which implies that the ATP concentration may affect the probability of ADP occupying the module. In such a scenario $\gamma$ would be dependent on the ATP concentration and a more elaborate model would be necessary to explain the dependence of dynein stepping on ATP concentration. Even if $\gamma$ does depend on ATP concentration, we expect its value to remain low even at high ATP concentrations since dynein exhibits high processivity. Even at low ATP concentrations where we expect $\gamma$ to decrease, most properties of the motor would not vary significantly from our predictions since the velocity, mean dwell time, mean step size, and probability of consecutive stepping do not change significantly at low $\gamma$ values. The exception to this is the run length and therefore, our estimation of the run length at low ATP concentration may be an underestimate of the actual value.

### 3.5.2 Constraints on the rates of the power stroke and ATP hydrolysis

There is a wide range of values in the literature for the rate of ATP hydrolysis/phosphate release[18,40,43,46]. We chose for our model the highest value of hydrolysis rate in the range, however, other values can be chosen as well. Since at

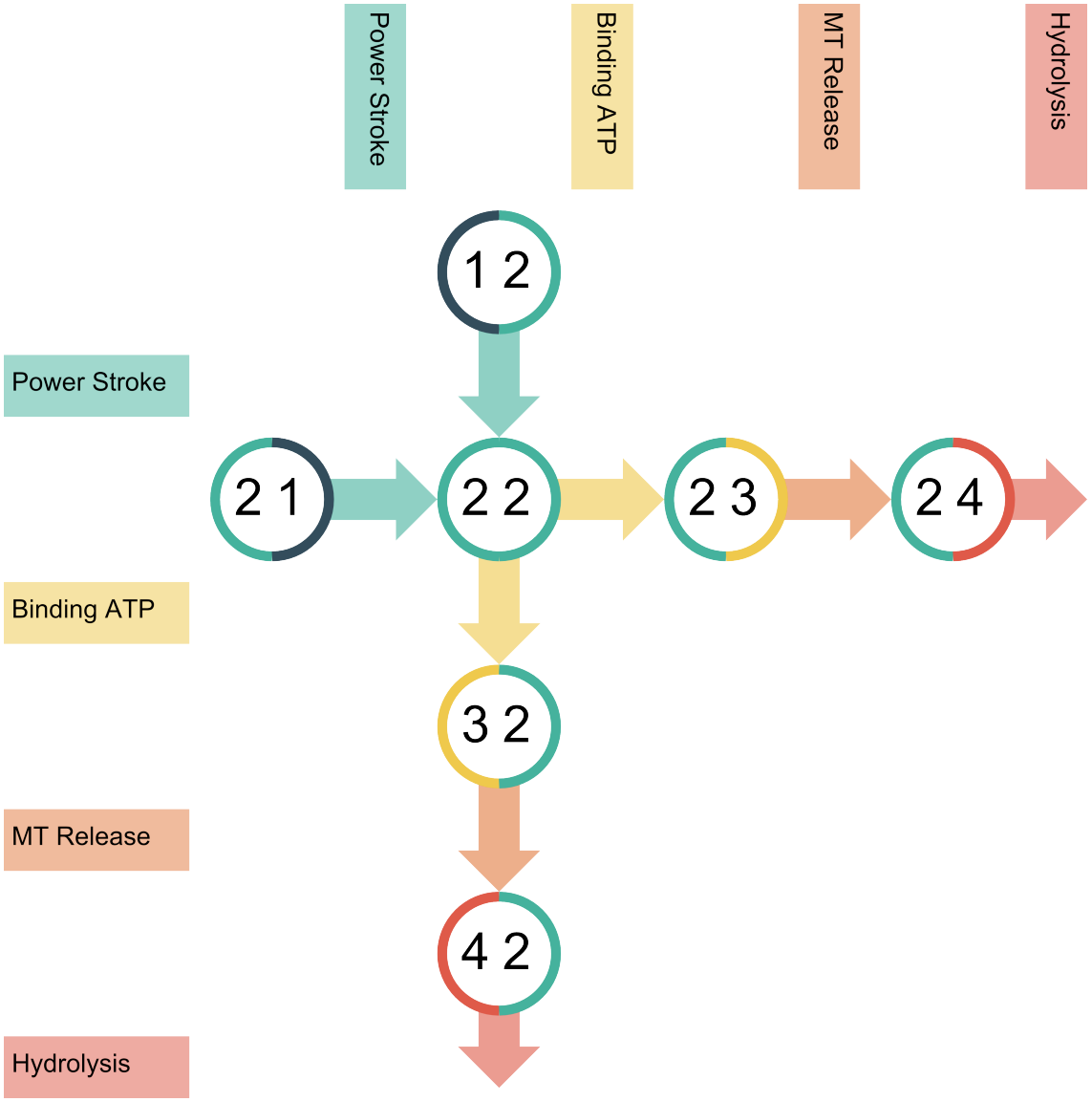


Figure 3.11: The kinetic scheme for a naive model for high ATP concentration conditions. In this model we assume that once one of the MDs binds ATP, the step is going to be carried out by that MD. The cycle stops for the partner MD until the step is complete. This of course cannot be technically true but it turns out to be a good approximation.

high ATP concentrations both ATP hydrolysis and the power stroke can be rate limiting, there exists a constraint on their values imposed by the velocity of the motor. The expression for the velocity as a function of the rates in our model is not simple and we therefore set out to find a simple approximation. We propose a naive model in which ATP hydrolysis occurs only once the power stroke is complete in both the MDs and once it occurs, only the hydrolyzing MD can resume its cycle and its partner MD waits until the step is complete (Fig. 3.11). In addition, rather than calculate the mean step size of a motor labeled dynein, we use the experimental value of the mean step size for a tail labeled motor and calculate the mean dwell time for the whole motor rather than a single MD. Even though this model is technically highly simplified, it approximates the velocity in our original model well as can be seen in Fig. 3.12. In this simplified model, the velocity can be expressed as:

$$V = \frac{\Delta x_{tail}}{\frac{1}{k_p} + \frac{1}{2k_{ATP}} + \frac{1}{k_d} + \frac{1}{k_h}} \tag{3.13}$$

where $\Delta x_{tail} = 6nm$.

We can define the constraint imposed on the hydrolysis and power stroke rates by the observed velocity:

$$k_h = \frac{2k_{ATP}k_pk_dV}{2k_{ATP}k_pk_d\Delta x_{tail} - 2k_{ATP}k_pV - 2k_{ATP}k_dV - k_pk_dV} \tag{3.14}$$

Furthermore, for a given velocity we can calculate the allowed minimum of each rate:

$$k_h^{min} = \frac{2k_{ATP}k_dV}{2k_{ATP}k_d\Delta x_{tail} - 2k_{ATP}V - k_dV} \tag{3.15}$$

$$k_p^{min} = \frac{2k_{ATP}k_dV}{2k_{ATP}k_d\Delta x_{tail} - 2k_{ATP}V - k_dV} \tag{3.16}$$

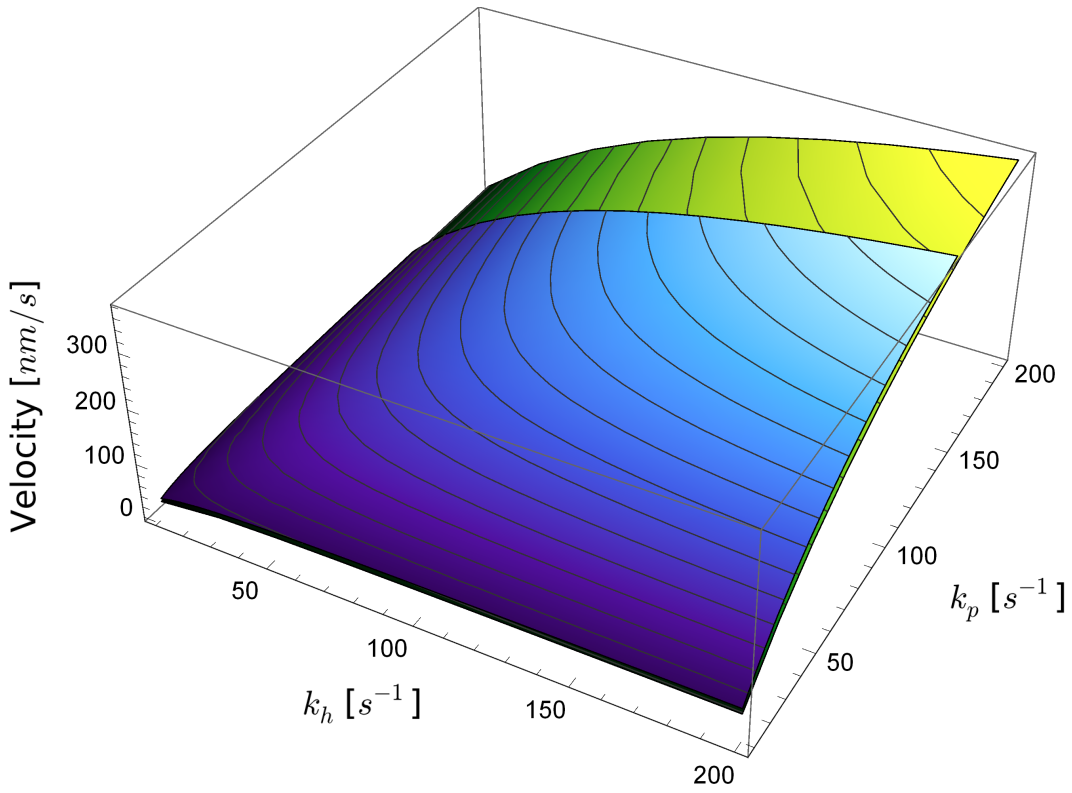


Figure 3.12: Comparison between the phase diagrams of the velocity as a function of the ATP hydrolysis and power-stroke rates. The blue surface is the surface of velocities obtained using our full unsimplified model. The green surface is the one obtained using our naive approach for high ATP concentration. As can be seen from the figure, the two surfaces are almost identical, justifying the use of our simplification.

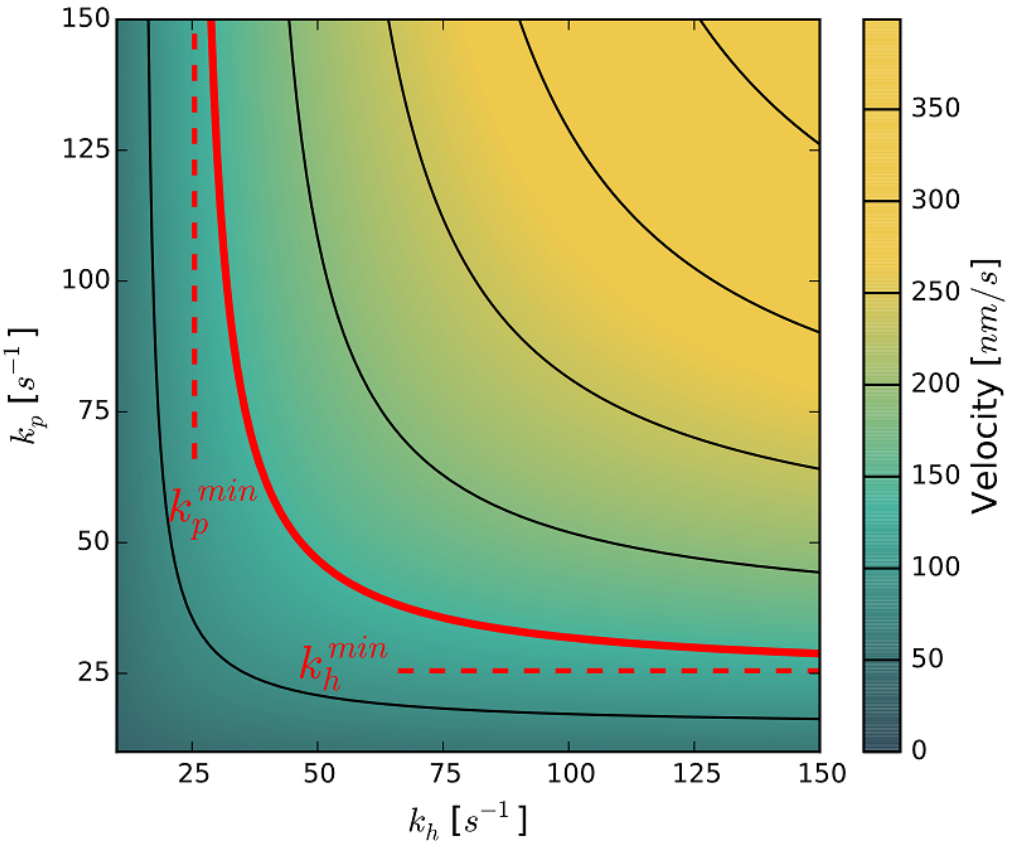


Figure 3.13: A phase diagram of the motor velocity as a function of the ATP hydrolysis rate, $k_h$ and the power-stroke rate, $k_p$. For a given velocity profile (such as the one represented by the red curve) the two rates become dependent and we can define lower bounds on their values (as represented by the dashed lines). Given the motor velocity and the value of one of the rates, one should be able to determine the remaining unknown value.

Our results, shown in Fig. 3.13, can help researchers determine the plausibility of experimentally measured rates.

## 3.6 Conclusions

We have constructed an analytically solvable model for the stepping mechanism of the dynein motor. Our approach, which consists of a coupling between a mechanical model and a kinetic scheme, can be used to reproduce a wide variety of properties of

the dynein motor such as step size and dwell time distributions, velocities, and run lengths. In addition we are able to reproduce 2 dimensional features of the dynein stepping pattern such as the inter motor orientation. Further more, our model allows us to make predictions as to how changes in ATP concentration and gating efficiency affect the behavior of the motor. Finally, using a simple Markov model, we are able to put constraint on the rates of the power stroke and ATP hydrolysis where there's uncertainty in the experimental results. Our model provides a more comprehensive picture of the function of the dynein motor and may be used successfully for other molecular motors. Additionally, we conclude that further investigation is needed in order to better understand the underlying molecular mechanisms behind the gating mechanism that account for motor coordination.

# Chapter 4

# Coarse Grained Simulations of Allostery in Dynein

## 4.1 Introduction

Dynein's function depends on its ability to translate the binding of ATP and MT into mechanical work. Binding of ATP leads to rapid dissociation from MT and to a priming stroke of the LN domain[41,46]. Binding to MT, on the other hand, accelerates ADP release and the power stroke[49]. Such a process is common in motor proteins. What makes dynein so interesting is the fact that the allosteric sites are located far from each other (see Fig. 2.1 in chapter 2). The dynein MTBD is located at the end of the stalk, an approximately 10nm long coiled coil structure. The main ATP binding site, on the other hand, is located at the opposite side of the AAA+ ring. Thus, the distance between both sites is about a 20-25nm. In order for the ATP binding site and the MTBD to communicate, an elaborate set of conformational changes has to occur[17,19,52]. Perhaps, the most noticeable conformational change is a relative shift in the position of the coils in the stalk[51,52,53].

The structural picture of how these allosteric transitions take place has become

clearer over the years, thanks to the availability of structures in a variety of allosteric states. However, there is still much we do not understand regarding the regulation of the motor's activity. ATP binding to the AAA3 unit and the application of an external load to the LN can affect and even block the allosteric communication between the ATP binding pocket and the MTBD[13,43,44,45,47]. This kind of regulation of dynein activity is particularly important in the context of gating.

The main role of gating is to increase the stepping probability of the trailing MD relative to that of the leading MD, which is needed to ensure that the motor moves forward and that it does so efficiently. Studies show that when a backward load is applied to the MD, it is less likely to detach from the MT than when the load is applied in the forward direction[47,58]. This suggests that the strain between the two MDs is responsible for gating.

In the context of dynein, gating could play an additional role. While both the MDs are bound to the MT, the likelihood of the motor falling off the MT completely is small. During a step, however, the chances of detachment increase dramatically as one of the two MDs is already detached. Since binding of ATP causes dynein to detach from the MT, the higher the ATP concentration, the more likely the motor is to fall off the MT. At biological ATP concentration, this probability is very high. Dynein, however, is a processive motor, taking many steps before detaching from the MT. This suggests that dynein may have a mechanism that prevents it from detaching from the MT during a step. While both MDs are bound to the MT there is some level of strain in the system. During a step, however, one of the MD is free with no net force acting on the bound MD. It is therefore likely that dynein detachment is tension dependent.

While the mechanism of force regulation in dynein is not entirely understood, recent studies indicate that the AAA3 unit may be involved in the process[43,43,44,45]. While the AAA3 units occupied by ADP, dynein activity is normal. When the

AAA3 unit is occupied by ATP, however, dynein is in a repressed state in which it is tightly bound to the MT regardless of ATP binding[43]. Furthermore, when tension is applied to the motor, MT detachment is possible once more[45]. This is consistent with the second role of the gating mechanism suggested above.

In order to gain a better understanding of the molecular basis of allosteric communication and gating in dynein, we performed Coarse Grained (CG) simulations of the MD. More specifically, we are interested in answering the following questions. What is the molecular mechanism underlying ATP/MT binding induced priming and power stroke? How does the AAA3 unit modulate the allosteric communication in the motor? What is the molecular basis of gating in dynein? The presented simulations have helped us gain insights into these questions.

## 4.2 Results

### 4.2.1 Allosteric Communication Between the ATP Binding Site and the MT

Studies show that binding of ATP to the primary ATP binding site between the AAA1 and AAA2 domains (referred to as AAA1/2) in dynein induces a conformational change in the LN domain (priming stroke) as well as unbinding from the MT[8,46,49]. On the other hand, binding of dynein to MT accelerates the release of ADP and the reverse conformational transition in the LN (power stroke)[41]. We require that our model reproduce this response of dynein to ATP and MT binding.

It is well established that allosteric communication between the AAA1/2 ATP binding site and the MTBD is transmitted through a sliding motion in the coiled coil stalk domain[50,51,52]. How the information is transfered between AAA1/2 and the stalk is less clear. However, it is reasonable to argue that the allosteric commu-

nication occurs via conformational changes in the AAA5 and AAA6 domains. To support this argument we rely on two pieces of evidence. First, structural studies of dynein suggest that the strut, a domain that connects the AAA5 and stalk domains, is involved in the conformational changes in the stalk[53]. Second, when comparing the crystal structures of dynein in the ATP bound and no nucleotide (apo) states, setting aside the LN and stalk domains, most of the conformational changes occur in the AAA1, AAA5, and AAA6 domains[17,18,19]. We therefore make the assumption, when constructing our model, that the allosteric communication pathway between AAA1/2 and the MTBD consists of the AAA5, AAA6, and the stalk. We refer to these domains as a group by AAA5/6/S.

We performed Brownian dynamics simulations, using a variant of the Self Organized Polymer (SOP) model with double well potentials (see appendix B) to monitor the conformational transitions that occur in response to either ATP or MT binding[68,69]. In order to make sure our model is consistent with experimental observations we simulated two scenarios. In scenario I, starting at the apo state, ATP binds to the AAA1 domain, inducing a conformational change along the AAA5/6/S allosteric pathway, leading to detachment from the MT. At the same time, the LN undergoes a priming stroke in reaction to the binding of the ATP molecule. In scenario II, the motor is in the ADP bound, pre power stroke state. Dynein then binds MT at the MTBD, accelerating the release of ADP as well as the power stroke in the LN domain.

In order to simulate the binding of ATP in scenario I we switched the Hamiltonian (see appendix B) of the AAA1/2 ATP binding domain from the apo state (referred to as state E) to the ATP bound state (referred to as state A). The most notable conformational change at the AAA1/2 site is the closure of the cleft between the AAA1 and AAA2 domains (Fig. 4.1). In order to track this conformational change we monitored $\Delta_{N1714-S2065}$, the distance between residue N1714 in AAA1

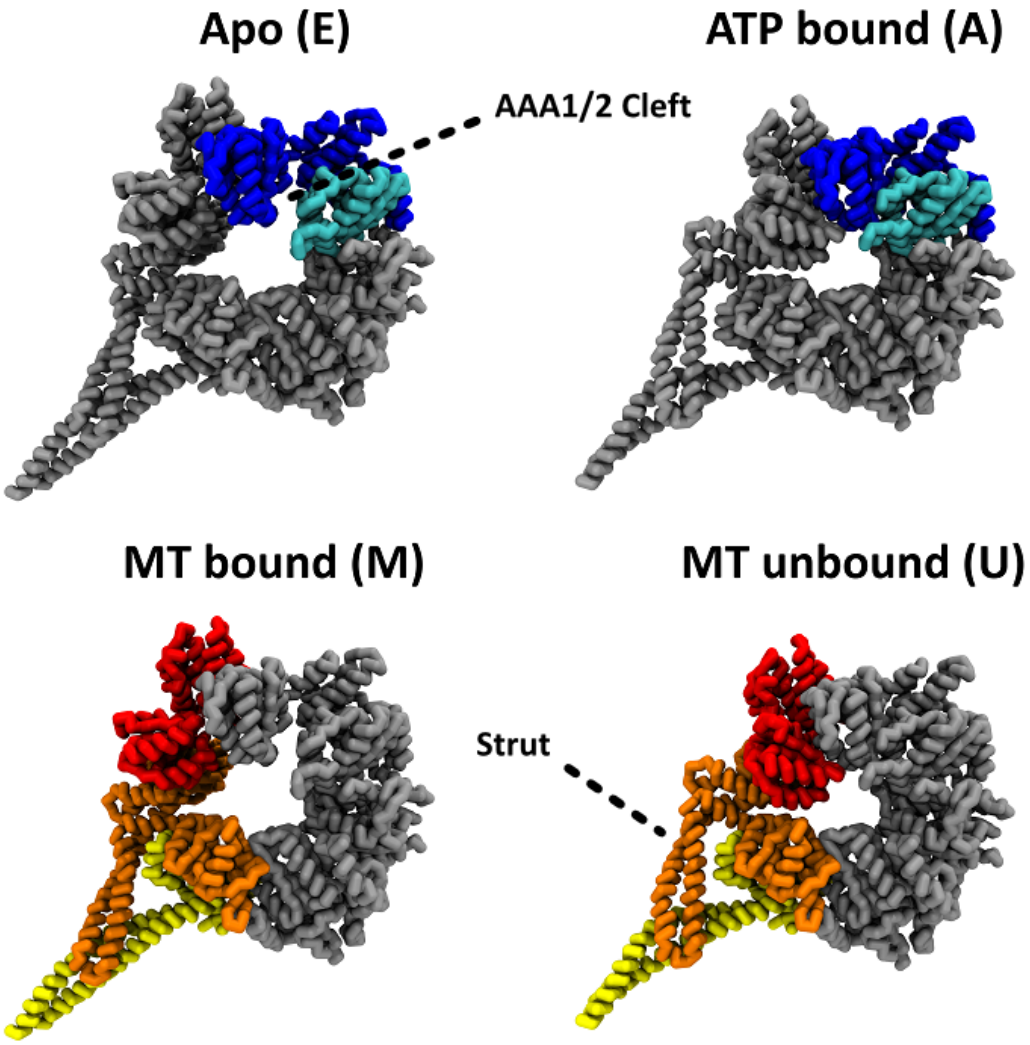


Figure 4.1: Conformational transitions in dynein. Top: Structures of dynein in the apo (E) state (left) or ATP bound (A) state (right). The AAA1 and AAA2 units are marked in blue and cyan respectively. The cleft between the AAA1 and AAA2 domains is open in the E state and closed in the A state. Bottom: Structures of dynein in the MT bound (M) state (left) and MT unbound (U) state (right). The AAA5, AAA6, and stalk domains are highlighted in orange, red, and yellow respectively. Notice the conformational differences between the two states, particularly in the strut sub-domain.

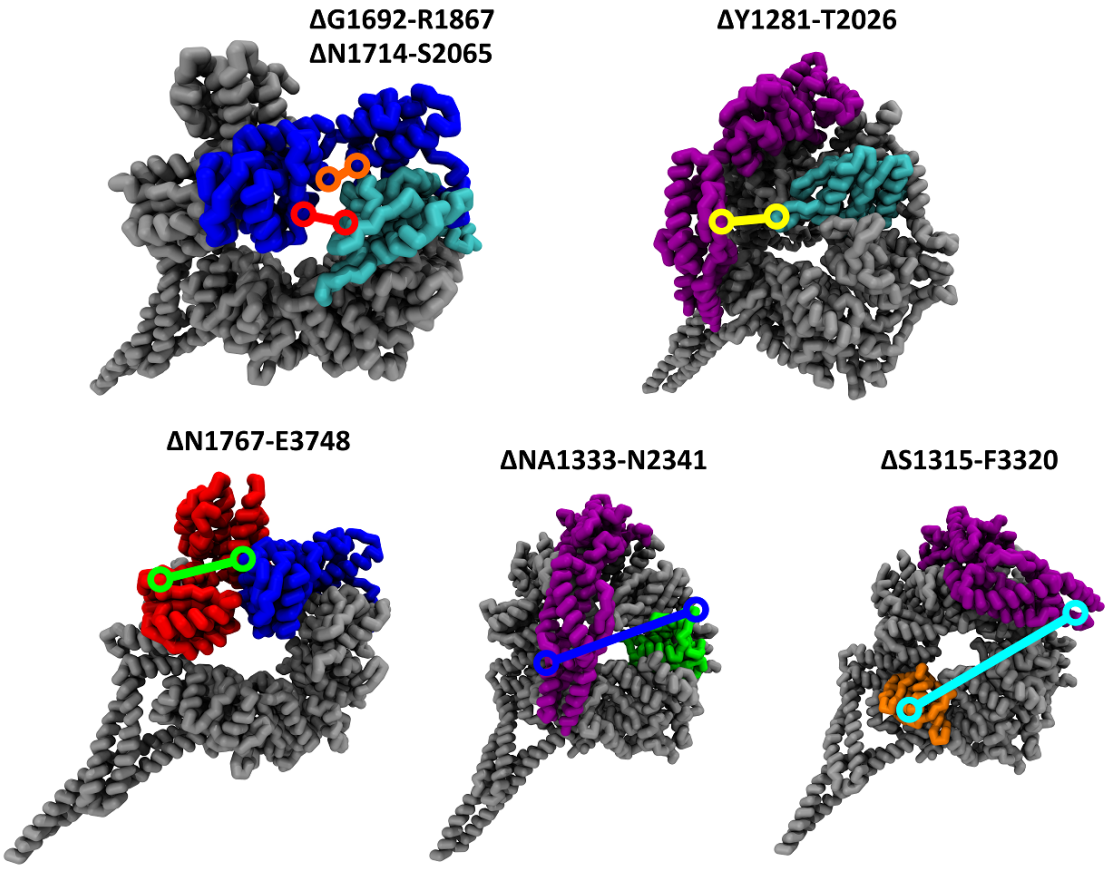


Figure 4.2: List of observables. The AAA1/2 ATP binding pocket is monitored by $\Delta_{G1692-R1867}$ (orange) while $\Delta_{N1714-S2065}$ (red) measures whether the AAA1/2 cleft is opened or closed. $\Delta_{Y1281-T2026}$ (yellow) measures the distance between the AAA2 ILs and the LN domain. The conformational transitions in the AAA5/6/S domains are monitored using $\Delta_{N1767-E3748}$ (green). $\Delta_{A1333-N2341}$ (blue) and $\Delta_{S1315-F3320}$ (cyan) measure the distance from the LN domain to the AAA3 and AAA5 domains respectively.

and residue S2065 in AAA2 (Fig. 4.2 and 4.3a). In order to account for the decreased affinity for MT, due to ATP binding, we switched the Hamiltonian for the AAA5/6/S domains from the MT bound state (referred to as state M) to the unbound state (referred to as state U). However, the switch between states M and U was made only after the distance $\Delta_{N1714-S2065}$ dropped below $8\mathring{A}$ in order to account for the fact that the allosteric transition occurs as a result of the conformational changes at the AAA1/2 site. We tracked the allosteric transition by monitoring $\Delta_{N1767-E3748}$, the distance between residues N1767 and E3748 (Fig. 4.2 and 4.3a).

In order to make sure that these transitions are a direct result of ATP binding and not a by product of the way we set up our model, we performed simulations of dynein in the apo state. In these simulations the Hamiltonian of the AAA1/2 was restrained to be in the E state. The Hamiltonian of the AAA5/6/S domains was still allowed to switch if $\Delta_{N1714-S2065}$ became smaller than $8\mathring{A}$. Fig. 4.3a shows that in the simulations of dynein with ATP the AAA1/2 site changed conformations, inducing a conformational change in AAA5/6/S. On the other hand, in the simulations of the apo state, no such transitions occurred. This is consistent with experimental results.

In scenario II, dynein binds to MT at the MTBD. In our simulations we do not include the MTBD explicitly but we assume that binding to the MT stabilizes the M state at the AAA5/6/S domains. Therefore, in order to simulate the effects of the MT binding, we switch the Hamiltonian at the AAA5/6/S domains from the U state to the M state. The AAA1/2 Hamiltonian is in the A state at the beginning of the simulations but the stabilizing interactions at the site are weakened (see appendix B) in order to reflect that the binding site contains ADP and not ATP. If $\Delta_{N1714-S2065}$ increases beyond 11Å, the Hamiltonian is switched to the E state which is nucleotide free.

Comparison between the M and U states show that binding of MT significantly

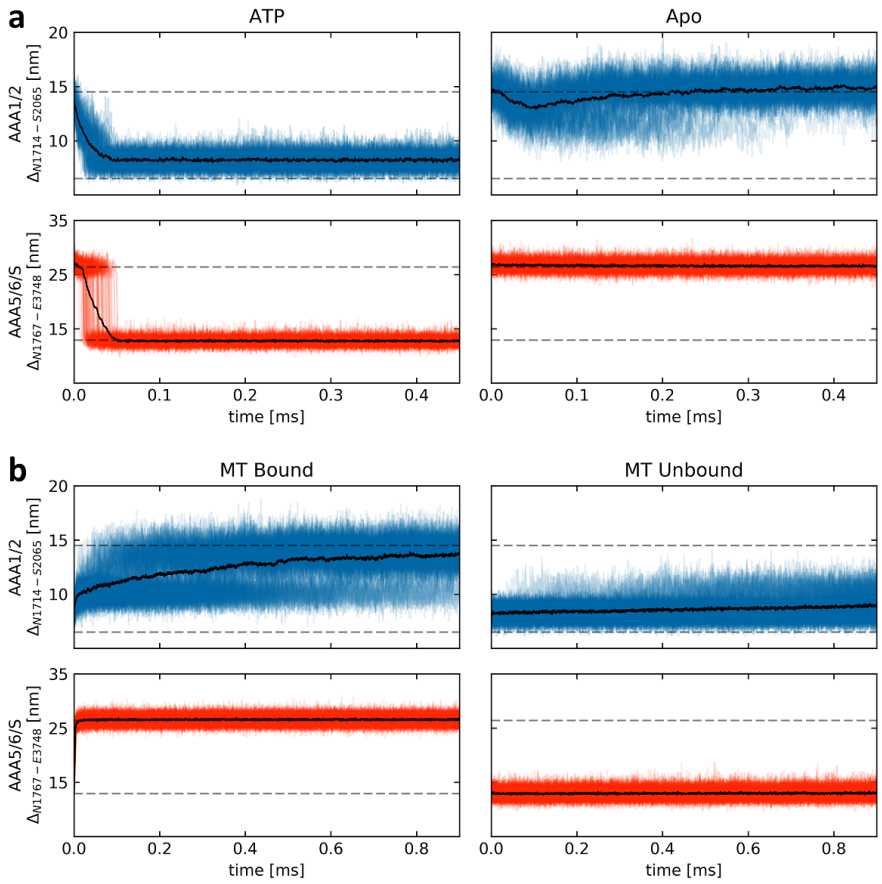


Figure 4.3: Dynein interactions with ATP and MT. a) Dynein response to ATP binding in 100 trajectories as a function of time. The values of $\Delta_{N1714-S2065}$ and $\Delta_{N1767-E3748}$ in the individual trajectories are plotted in blue and red respectively. The black curves represent the ensemble average values. Dashed lines mark the values of $\Delta_{N1714-S2065}$ and $\Delta_{N1767-E3748}$ in each state in the reference structures. In the presence of ATP the AAA1/2 cleft closes rapidly (upper left). The AAA5/6/S domains transition from the M state to the U state in response to the AAA1/2 transition (lower left). In the absence of ATP, the AAA1/2 cleft remains open (upper right) and there are no transitions in the AAA5/6/S domains (lower right). b) Dynein response to MT binding. In the absence of MT, the AAA5/6/S domains remain in the U state (lower right). The AAA1/2 cleft appears to show signs of opening, however, this process is slow and we do not observe full opening of the cleft (upper right). Once the MT binds to dynein the AAA5/6/S domains transition to the M state (lower left). This dramatically accelerates the opening of the AAA1/2 cleft (upper left).

accelerates the opening of the cleft at the AAA1/2 site (Fig. 4.3b). Furthermore, while there are fluctuations in the AAA1/2 cleft in the simulations without MT, the cleft does not open fully within the simulation time. While $\Delta_{N1714-S2065}$ is a good measure of the opening of the AAA1/2 cleft, it is not necessarily a good indicator of what happens at the ATP binding site itself. In order to monitor the conformational changes at the ATP binding site directly, we measured $\Delta_{G1692-R1867}$, the distance between residues G1692 and R1867, located at ATP binding site (Fig. 4.2). We tracked $\Delta_{N1714-S2065}$ and $\Delta_{G1692-R1867}$ simultaneously, both at the M and U states. Fig. 4.4a shows that in the absence of MT, while there is partial opening of the AAA1/2 cleft, the ATP binding site itself remains closed. In contrast, in the simulations in which the AAA5/6/S domains adopted the M state, both the AAA1/2 cleft and ATP binding site are open (Fig. 4.4b). This supports the hypothesis that conformational changes along the AAA5/6/S pathway are responsible for the acceleration of ADP release from the AAA1/2 site.

### 4.2.2 Conformational Transitions in the LN Domain

In addition to regulation of dynein's affinity for MT, ATP binding and hydrolysis controls the conformations of the LN domain. More specifically, binding of ATP at the AAA1/2 site leads to a priming stroke in the LN domain. Binding of MT, on the other hand, accelerates a power stroke motion in the LN domain[49]. To test whether our simulations are consistent with these findings we tracked the motion of the LN domain in both scenarios I and II.

In scenario I the system is found in the post stroke apo state. As we have already discussed, binding of ATP to the AAA1/2 site leads to a reduced affinity for MT. In order to track the response of the LN domain to the binding of ATP, we measured $\Delta_{S1315-F3320}$, the distance between residue S1315 near the N-terminus of the LN domain and residue F3320 in the AAA5 domain (Fig. 4.2). Both residues

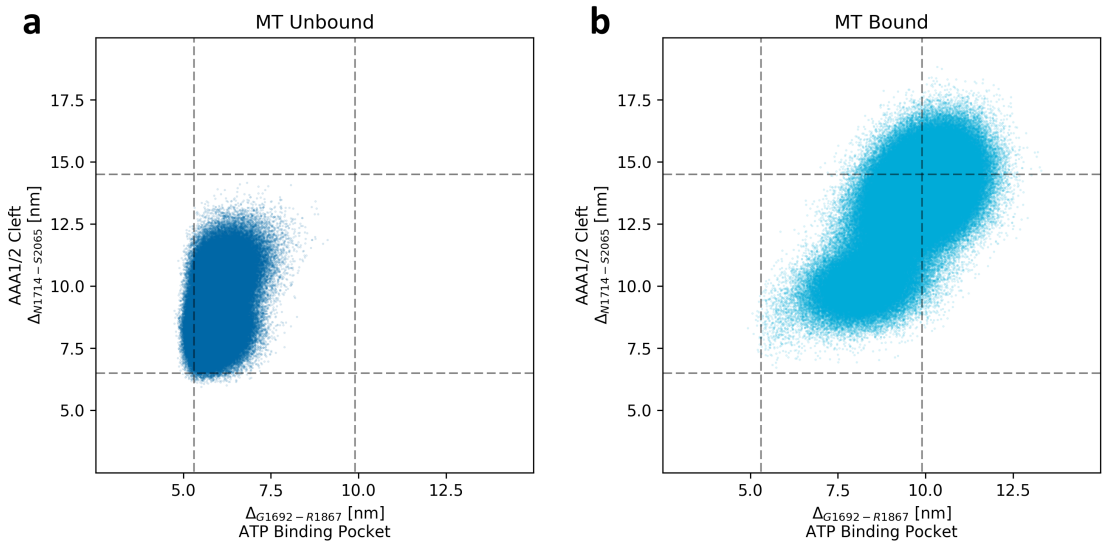


Figure 4.4: Conformational transitions of the AAA1/2 domains. a) Scatter plot of $\Delta_{N1714-S2065}$ (AAA1/2 cleft) against $\Delta_{G1692-R1867}$ (ATP binding pocket) in 100 trajectories of dynein in the absence of MT (blue). While the AAA1/2 cleft partially opens, the ATP binding pocket remains predominantly closed. b) Scatter plot of $\Delta_{N1714-S2065}$ (AAA1/2 cleft) against $\Delta_{G1692-R1867}$ (ATP binding pocket) in 100 trajectories of dynein when bound to MT (cyan). Both AAA1/2 cleft and ATP binding pocket are predominantly open. The $\Delta_{N1714-S2065}$ and $\Delta_{G1692-R1867}$ values in the reference structures are marked by dashed lines.

are found in the interface between the LN and the AAA5 domains, which makes $\Delta_{S1315-F3320}$ a good measure of whether the LN domain is in the post stroke conformation or whether the LN domain is detached from the AAA5 binding site.

Fig. 4.5 shows that in a large percentage of our simulations the LN domain detaches from the AAA5 domain. To make sure that this is a result of ATP binding we monitored $\Delta_{S1315-F3320}$ in simulations of the apo state with no ATP bound at the AAA1/2 site. With the exception of a couple of trajectories, the LN domain remained bound to the AAA5 domain, as expected. In order to find out whether the LN domain reaches the AAA3 binding site in the pre stroke state, we measured $\Delta_{A1333-N2341}$, the distance between residue A1333 in the LN domain and residue N2341 in the AAA3 domain. These residues come into contact in the pre stroke state and therefore $\Delta_{A1333-N2341}$ is an indicator of whether the priming stroke is complete. Fig. 4.5 shows that in several trajectories the LN domain reaches its pre stroke conformation within the simulation time frame.

Similarly, in scenario II we expect the binding of MT to accelerate the power stroke. Fig. 4.6 shows that the LN-AAA3 detachment time scale is significantly shorter when MT is bound to dynein. Furthermore, Fig. 4.6 shows that in several trajectories, in which MT is bound to dynein, the LN domain reaches its AAA5 binding site, thus completing the power stroke. Thus, we conclude that our model reflects properly the fact that MT binding accelerates both ADP release and the power stroke.

The mechanism with which conformational changes in the AAA1/2 cleft affect the stability of the pre stroke state is evident. The segment of the LN that connects the AAA3 domain to the AAA1 domain is almost perpendicular to the AAA1/2 cleft. This means that any opening motion in the cleft would pull away the LN domain from its binding site in the AAA3 domain, thus destabilizing the pre stroke state. We plotted $\Delta_{A1333-N2341}$ (LN-AAA3 distance) against $\Delta_{N1714-S2065}$ (AAA1/2

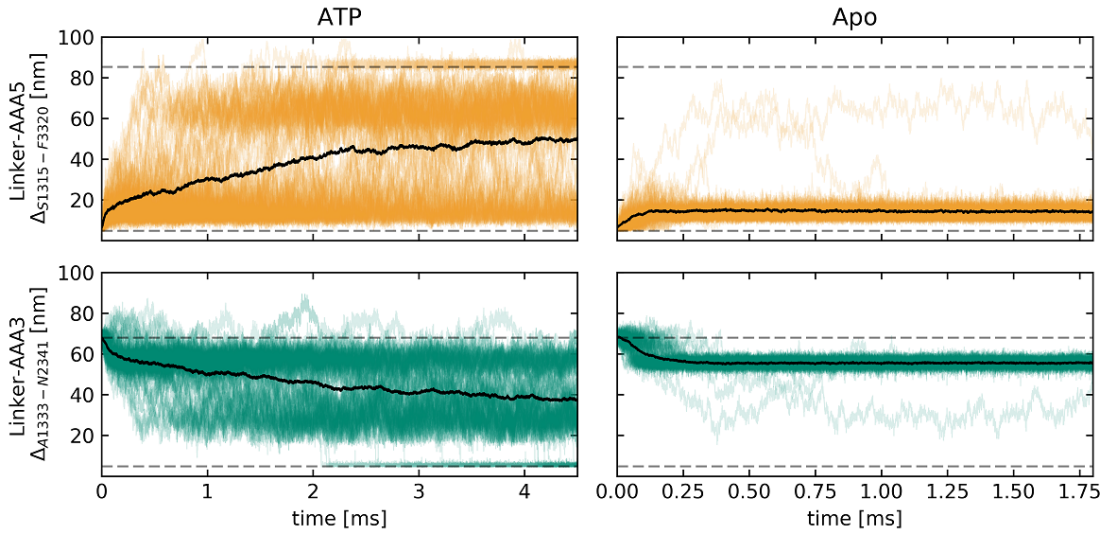


Figure 4.5: The dynein priming stroke. The time response of the LN to ATP binding is tracked using $\Delta_{S1315-F3320}$ (LN-AAA5 distance) and $\Delta_{A1333-N2341}$ (LN-AAA3 distance) in 100 trajectories. Individual trajectories are plotted in orange and green respectively. The ensemble averages are represented by the black curves. The $\Delta_{S1315-F3320}$ and $\Delta_{A1333-N2341}$ values in the reference structures are marked by dashed lines. When dynein binds ATP the LN detaches from the AAA5 domain (upper left). The distance between the LN and the AAA3 decreases until the LN reaches the pre stroke conformation (lower left). In the absence of ATP, the LN remains bound to the AAA5 domain (upper right) and far from the AAA3 unit (lower right).

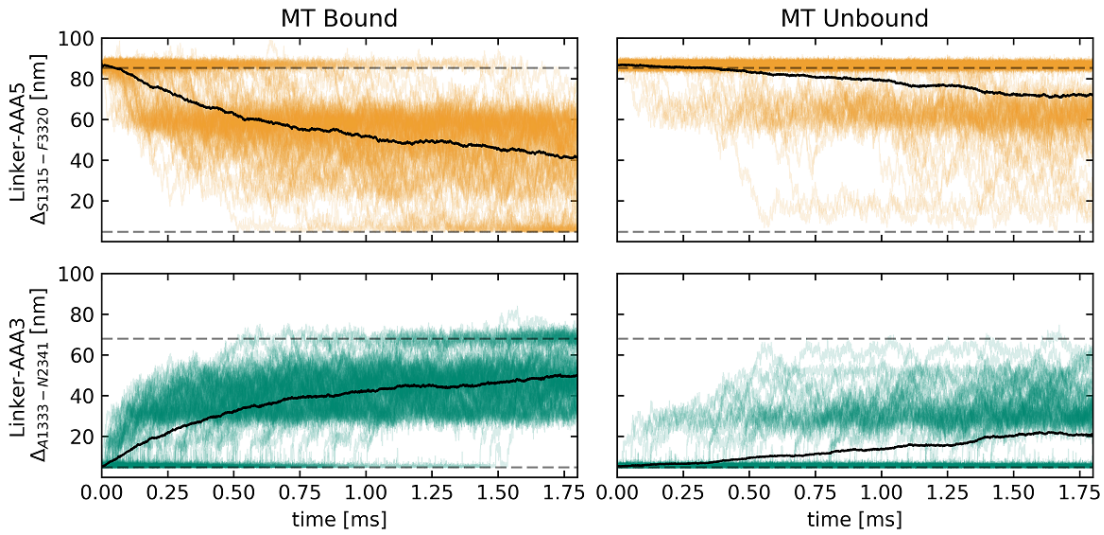


Figure 4.6: The dynein power stroke. The time response of the LN to MT binding is tracked, similarly to Fig. 4.5. When dynein is bound to MT the LN's detachment rate from the AAA3 domain (lower left) is faster when compared to the detachment in the absence of MT (lower right). The distance of the LN from the AAA5 binding site decreases until binding occurs, completing the power stroke. The rate of the power stroke is noticeably faster when dynein is bound to MT (upper left) in comparison to unbound dynein (upper right). The $\Delta_{S1315-F3320}$ and $\Delta_{A1333-N2341}$ values in the reference structures are marked by dashed lines.

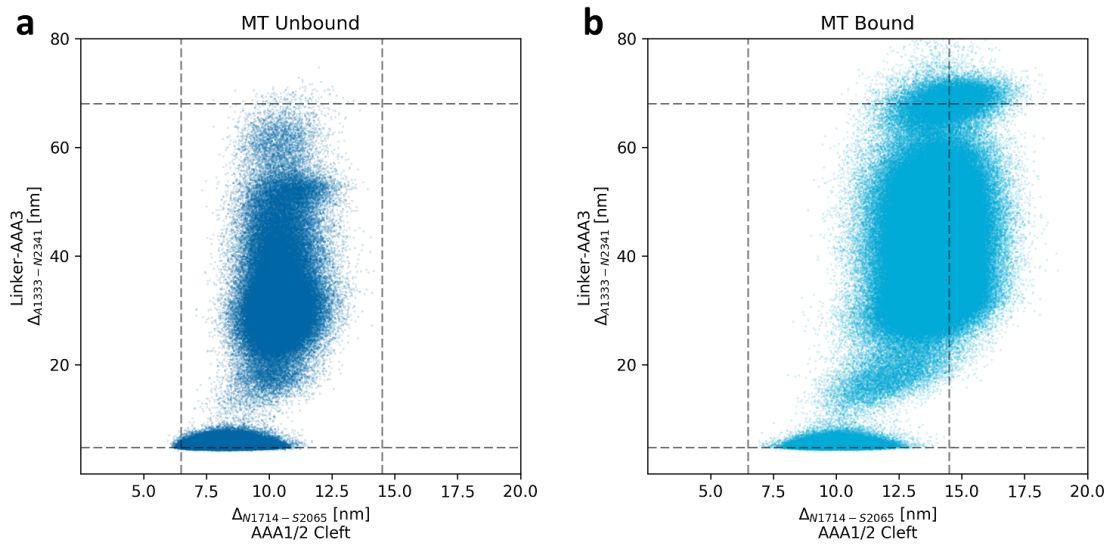


Figure 4.7: Mechanism of LN-AAA3 detachment. a) Scatter plot of $\Delta_{A1333-N2341}$ (LN-AAA3 distance) against $\Delta_{N1714-S2065}$ (AAA1/2 cleft) in 100 trajectories of dynein in the absence of MT (blue). When the AAA1/2 cleft is closed, the LN remains mostly bound to the AAA3 binding site. b) Scatter plot of $\Delta_{A1333-N2341}$ (LN-AAA3 distance) against $\Delta_{N1714-S2065}$ (AAA1/2 cleft) in 100 trajectories of dynein bound to MT (cyan). As the AAA1/2 cleft opens, the LN tends to detach from the AAA3 domain and undergo a power stroke. The $\Delta_{A1333-N2341}$ and $\Delta_{N1714-S2065}$ values in the reference structures are marked by dashed lines.

cleft) to illustrate this. Fig. 4.7a and 4.7b show that as the AAA1/2 cleft tends to be more open, due to the binding of MT, the LN is more likely to detach from the AAA3 binding site. Similarly, closing the AAA1/2 cleft would increase the stability of the pre stroke conformation.

The picture is more complex when it comes to the post stroke state. Our simulations indicate that the LN regularly comes into contact with the AAA2 ILs in the post stroke state. In fact, the interaction between the LN and the AAA2 ILs seems to lead to slight detachment of the LN from the AAA5 binding site even in the apo state(Fig. 4.8b). This is not entirely surprising. In the crystal structure by Kon *et al* the LN domain is shown to interact with the IL in the ADP bound state (which is structurally similar to the apo state)[18]. The LN is also somewhat shifted from its AAA5 bound position in this structure, and is located

in the cleft between AAA4 and AAA5 domains. Naively, this may suggest that favorable interactions between the LN and the AAA2 domains are not compatible with LN-AAA5 docking. Nevertheless, these interactions seem to stabilize the post stroke conformation as a whole (Fig. 4.8a and 4.8b). To further support this claim, we ran mutation simulations of the apo state in which the stabilizing interactions between the LN and the AAA2 domain are switched off. Fig. 4.9 shows that without stabilizing interactions between the LN and the ILs, the post stroke conformation is destabilized as the LN is more likely to detach from the AAA5 domain.

The involvement of the AAA2 ILs in the stabilization of the post stroke state presents a possible explanation of how binding of ATP to the AAA1/2 binding site destabilizes this state. Closing of the AAA1/2 cleft bring the AAA2 ILs close to the AAA1 domain, making them unavailable to interact with N-terminus adjacent regions in the LN domain. This is illustrated by Fig. 4.10 in which we track the distance between the LN and the AAA2 ILs and the transition of the AAA1/2 cleft. While the AAA1/2 cleft remains open, the AAA2 ILs are found to be mostly in contact with the LN domain. When the cleft closes, however, due to the binding of ATP, these contacts are destabilized. This implies that the AAA2 ILs serve as a control switch for the post stroke state.

### 4.2.3 Inactivation of Dynein Through the AAA3 Domain

It is well established that the AAA1/2 domain is the principle site for ATP binding and hydrolysis in Dynein[70]. Nevertheless, several studies show that nucleotide binding and hydrolysis occurs at two additional sites, located in the AAA3 and AAA4 domains respectively. Furthermore, recent studies revealed that the AAA3 plays a crucial role in the regulation of dynein activity[43,44,45]. More specifically, while the AAA3 site contains ADP, dynein behaves normally. However, when the site is empty or is occupied by ATP, the LN domains appears to be locked in the

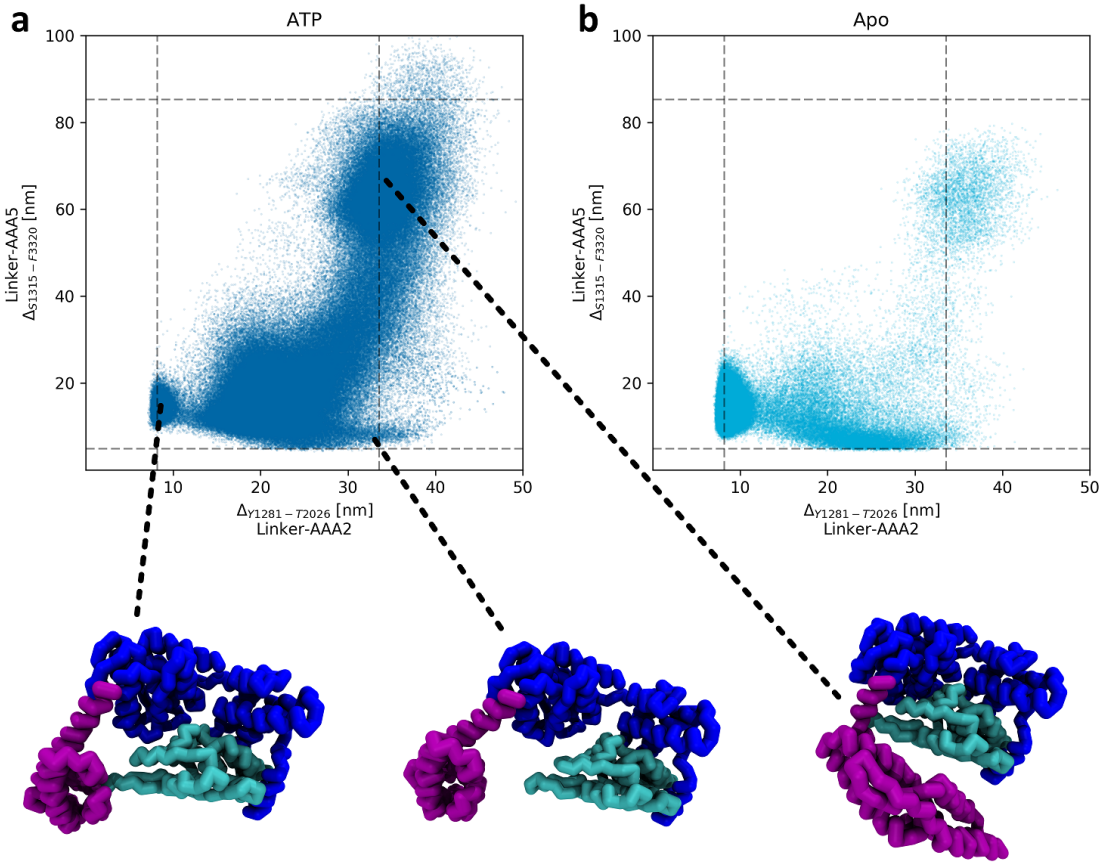


Figure 4.8: Post stroke stabilization by AAA2 ILs. a) Scatter plot of $\Delta_{S1315-F3320}$ (LN-AAA5 distance) against $\Delta_{Y1281-T2026}$ (LN-AAA2 distance) in 100 trajectories of ATP bound dynein (blue). Before the LN detaches from the AAA5 domain, the AAA2 ILs tend to form contacts with the LN. Eventually the LN detaches from both domains due to the closure of the AAA1/2 cleft. b) Scatter plot of $\Delta_{S1315-F3320}$ (LN-AAA5 distance) against $\Delta_{Y1281-T2026}$ (LN-AAA2 distance) in 100 trajectories of dynein in the apo state (cyan). In the absence of ATP, the LN still binds to the AAA2 ILs, suggesting that these interactions stabilize the post stroke conformations. The $\Delta_{S1315-F3320}$ and $\Delta_{Y1281-T2026}$ values in the reference structures are marked by dashed lines. The three structures in the bottom illustrate the AAA2 bound state (left), AAA5 bound state (middle), and pre stroke state (right).

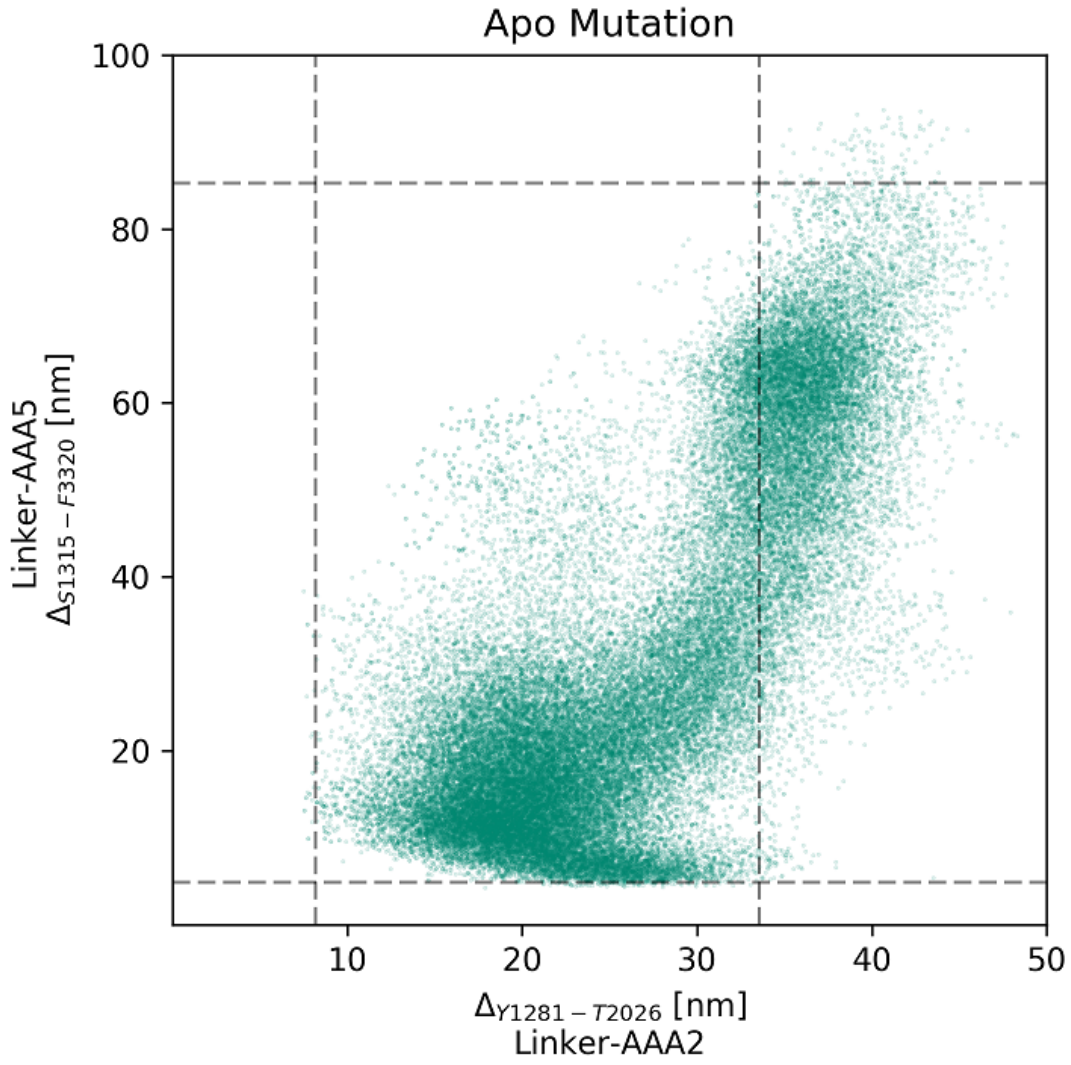


Figure 4.9: Post stroke destabilization by AAA2 mutation. Scatter plot of $\Delta_{S1315-F3320}$ (LN-AAA5 distance) against $\Delta_{Y1281-T2026}$ (LN-AAA2 distance) in 20 trajectories of dynein in the apo state with no stabilizing interactions between the LN and the AAA2 domain (geen). Without stabilization of the post stroke conformation by the LN-AAA2 interactions, the LN detaches from the AAA5 domain even in the absence of ATP. The $\Delta_{S1315-F3320}$ and $\Delta_{Y1281-T2026}$ values in the reference structures are marked by dashed lines.

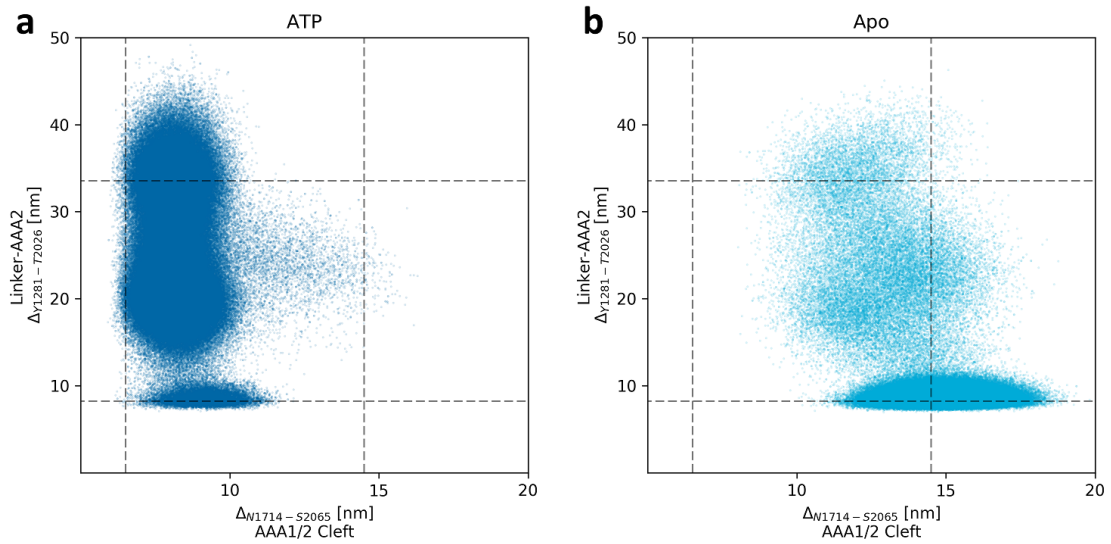


Figure 4.10: Mechanism of LN-AAA5 detachment. a) Scatter plot of $\Delta_{Y1281-T2026}$ (LN-AAA2 distance) against $\Delta_{N1714-S2065}$ (AAA1/2 cleft) in 100 trajectories of ATP bound dynein (blue). Closing of the AAA1/2 cleft leads to LN-AAA2 detachment as the ILs become unavailable for interaction with the LN. b) Scatter plot of $\Delta_{Y1281-T2026}$ (LN-AAA2 distance) against $\Delta_{N1714-S2065}$ (AAA1/2 cleft) in 100 trajectories of dynein in the apo state (cyan). When the AAA1/2 cleft remains open the AAA2 ILs are available to interact with the LN domain. The $\Delta_{Y1281-T2026}$ and $\Delta_{N1714-S2065}$ values in the reference structures are marked by dashed lines.

post stroke conformation regardless of the nucleotide state of the AAA1/2 site[43]. Further studies show that in this repressed state dynein maintains high affinity for MT as well[43,45].

In order to investigate the mechanism with which the AAA3 domain represses dynein activity we simulated scenario I using the crystal structure of the repressed system instead of the crystal structure of the apo state as reference for our Hamiltonian[43]. We monitored the motion of the LN domain in the simulations by tracking $\Delta_{S1315-F3320}$ as a function of time. In contrast to the non repressed system, Fig. 4.11 shows that in most of the trajectories, the LN domain remained bound to the AAA5 binding site despite the fact that the AAA1/2 Hamiltonian has been switched to the A state. This is consistent with the results of Bhabha *et al*[43]. We investigated the repression mechanism further by plotting $\Delta_{N1714-S2065}$ (AAA1/2 cleft) against $\Delta_{G1692-R1867}$ (AAA1/2 ATP binding site) (Fig. 4.12). Interestingly, the ATP binding site itself does close as a result of switching the Hamiltonian to the ATP bound reference structure. However, the AAA1/2 cleft remained open in the majority of trajectories.

We already pointed out the involvement of the AAA2 domain in the stabilization of the post stroke state. This is further supported by structural and mutational studies that show that the ILs play an important role in the regulation of dynein activity[18,43]. More specifically, the ILs come into contact with the LN domain in the repressed structure[18,43]. To better understand the involvement of the ILs in the repression mechanism we performed simulation of a mutated system in which the stabilizing interactions between the ILs and the LN where replaced with repulsive interactions. Interestingly, when these interactions where turned off, the AAA1/2 cleft was able to close and the LN was able to detach from the AAA5 domain, despite the fact that the AAA3 domain was in the ATP bound state, (Fig. 4.11 and 4.13). Therefore, our simulation support the hypothesis that the AAA2 ILs are

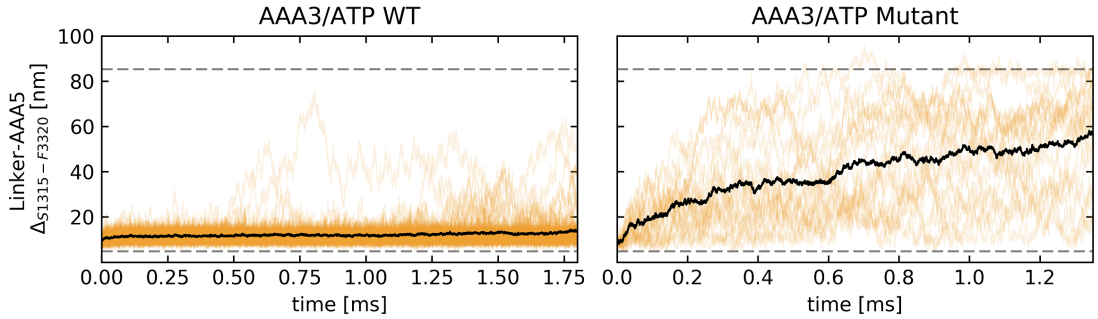


Figure 4.11: ATP bound to the AAA3 domain represses dynein activity. The time response of the LN to ATP binding in the repressed state is tracked using $\Delta_{S1315-F3320}$ (LN-AAA5 distance) and $\Delta_{A1333-N2341}$ (LN-AAA3 distance) in 100 trajectories. Individual trajectories are plotted in orange and the ensemble averages are represented by the black curves. The $\Delta_{S1315-F3320}$ and $\Delta_{A1333-N2341}$ values in the reference structures are marked by dashed lines. Even when the AAA1/2 site is occupied by ATP, the LN remains bound to the AAA5 domain in the repressed state (left). 20 trajectories of the mutated motor in which LN-AAA2 interactions are turned off show that the LN does detach from the AAA5 domain (right). This implies that the LN-AAA2 contacts are essential for the system repression.

crucial for the repression mechanism.

### 4.2.4 The AAA2 ILs (Insert Loops) Involvement in Motor Gating

In order to investigate how gating works in dynein, we performed simulations of dynein in the repressed state in scenario I, while applying an constantly increasing force at the LN N-terminus. We performed two sets of these simulations. In one set, the direction of the force was in the negative direction along the MT axis, and in the second set, the direction of the force was along the positive direction (Fig. 4.14). As the magnitude of the force increased, the AAA1/2 cleft closed, indicating that the allosteric communication between the AAA1/2 site and the AAA5/6/S pathway was no longer repressed (Fig. 4.15). Interestingly, detachment of the LN from the AAA2 ILs was not required for the AAA1/2 cleft to close (Fig. 4.16a and 4.16b). The criteria for the closing of the AAA1/2 cleft was detachment of the LN

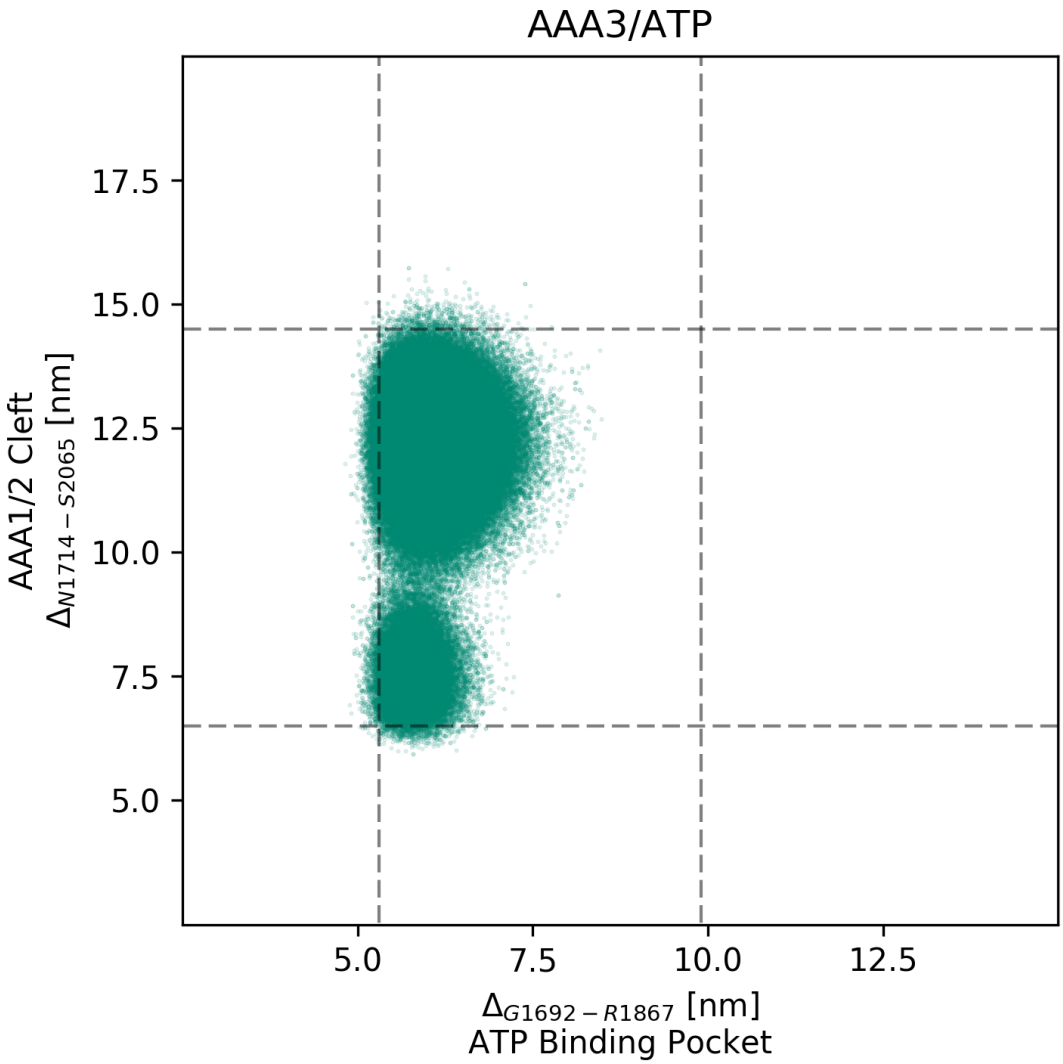


Figure 4.12: Scatter plot of $\Delta_{N1714-S2065}$ (AAA1/2 cleft) against $\Delta_{G1692-R1867}$ (ATP binding pocket) in 100 trajectories of ATP bound dynein in the repressed state (green). The $\Delta_{N1714-S2065}$ and $\Delta_{G1692-R1867}$ values in the reference structures are marked by dashed lines. In the repressed state the AAA1/2 cleft remains open despite the ATP binding pocket being occupied by ATP and closed. This is a result of the AAA2 ILs contacts with the LN.

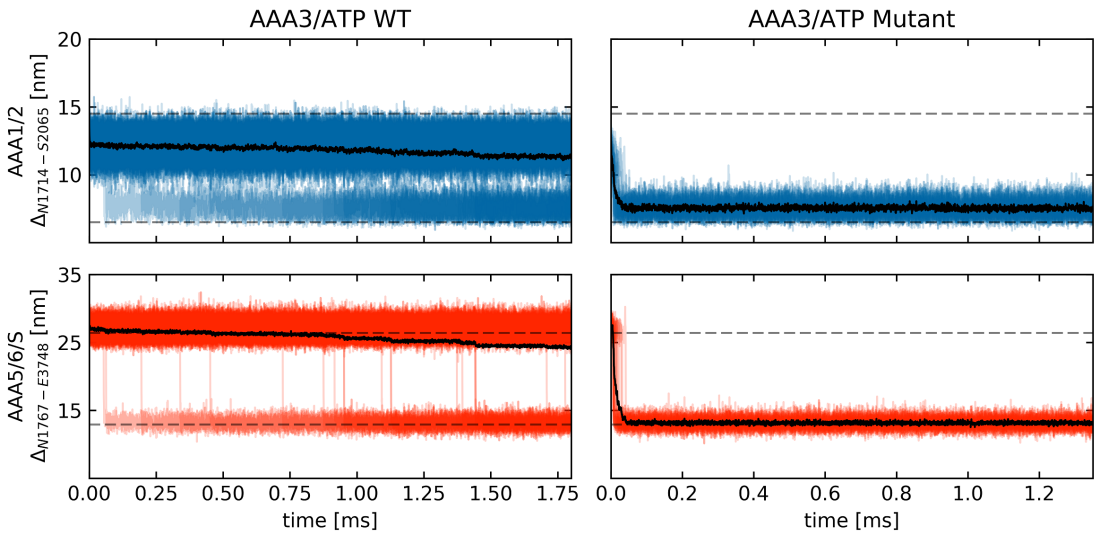


Figure 4.13: Allosteric inhibition in the repressed state. Dynein response to ATP binding in 100 trajectories as a function of time. The values of $\Delta_{N1714-S2065}$ and $\Delta_{N1767-E3748}$ in the individual trajectories are plotted in blue and red respectively. The black curves represent the ensemble average values. Dashed lines mark the values of $\Delta_{N1714-S2065}$ and $\Delta_{N1767-E3748}$ in each state in the reference structures. In the repressed state the ATP induced conformational transitions of the AAA1/2 cleft (upper left) and the AAA5/6/S domains (lower left) occur at a significantly slower rate when compared with the non repressed state (Fig. 3.3). In the 20 trajectories of the AAA2 mutation simulations the allosteric pathway operates normally as in the active state.

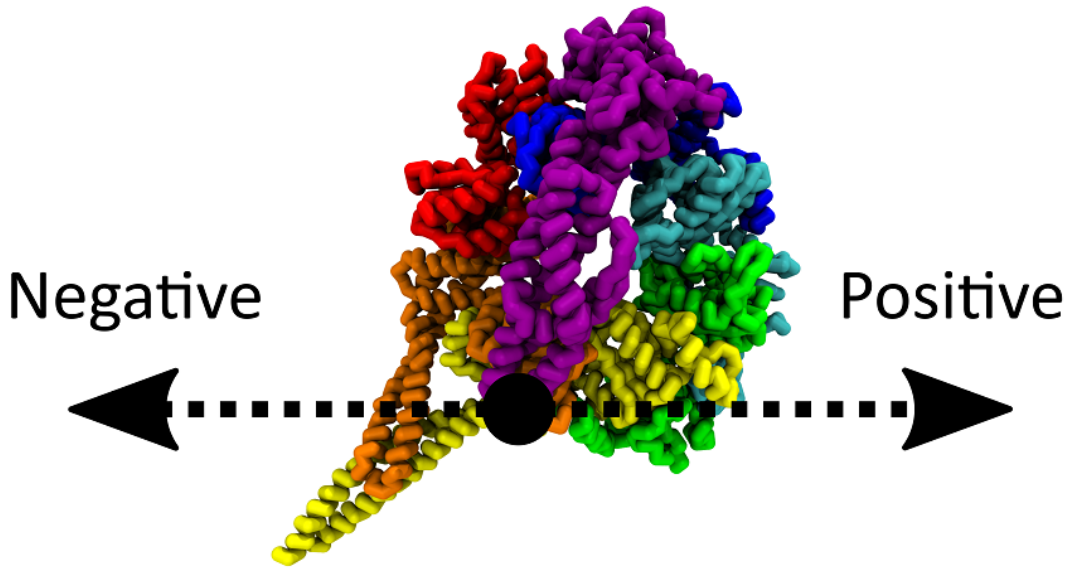


Figure 4.14: Direction of external force. The figure illustrates how the external force is applied at the N-terminus of the LN domain (black dot). The two arrows show the negative and positive directions of the force along the MT axis.

from the AAA5 binding site (Fig. 4.16c and 4.16d). This suggests that the AAA1/2 cleft can close even when the LN is bound to the ILs as long as it is not bound to the AAA5 domain as well. LN release from the AAA5 site allows it to assume conformations that can accommodate the closing of the AAA1/2, thus allowing the allosteric communication pathway to function normally.

We calculated the distribution of the force required to allow for the closure of the AAA1/2 cleft (Fig. 4.17). Closure of the AAA1/2 cleft leads to rapid detachment from the MT and, making the histogram in Fig. 4.17 an indicator of the unbinding force of the motor from the MT. In both simulation sets we obtained mean unbinding forces of the order of $\approx 40pN$, which is about an order of magnitude larger than the unbinding forces measured by the experimental set up[44,45,47]. This is likely due to the nature of the CG model. Nevertheless, we can still gain insights into the effects

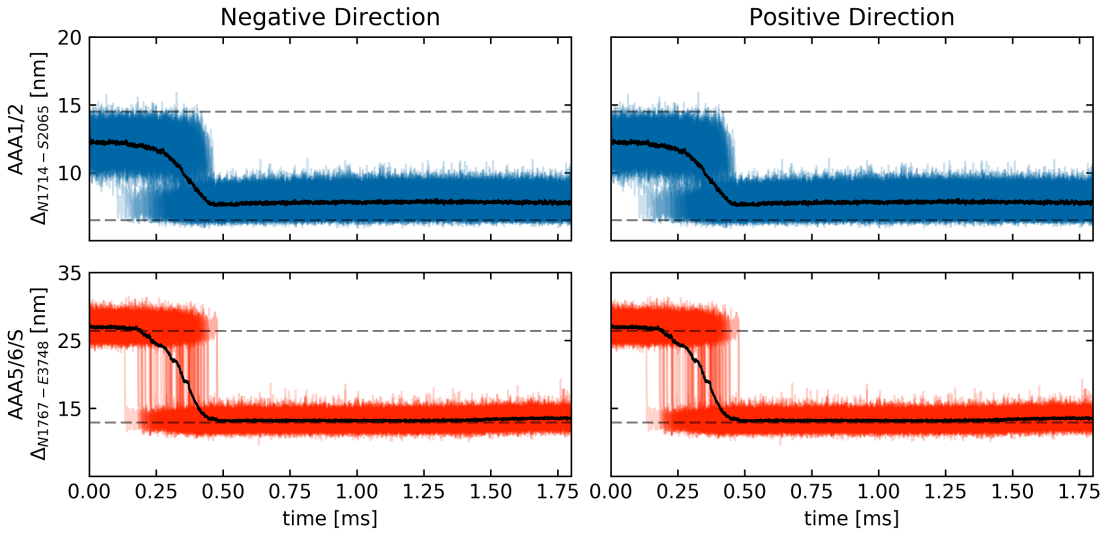


Figure 4.15: Applied force reactivates allosteric communication in dynein. Dynein's response to external force as a function of time in 100 trajectories of dynein in ATP bound repressed state. The values of $\Delta_{N1714-S2065}$ and $\Delta_{N1767-E3748}$ in the individual trajectories are plotted in blue and red respectively. The black curves represent the ensemble average values. Dashed lines mark the values of $\Delta_{N1714-S2065}$ and $\Delta_{N1767-E3748}$ in each state in the reference structures. When a large enough force is applied at the N-terminus of the LN in the negative (upper and lower left) and positive (upper and lower right) direction along the MT axis, the allosteric communication resumes normal activity despite dynein being in the repressed state.

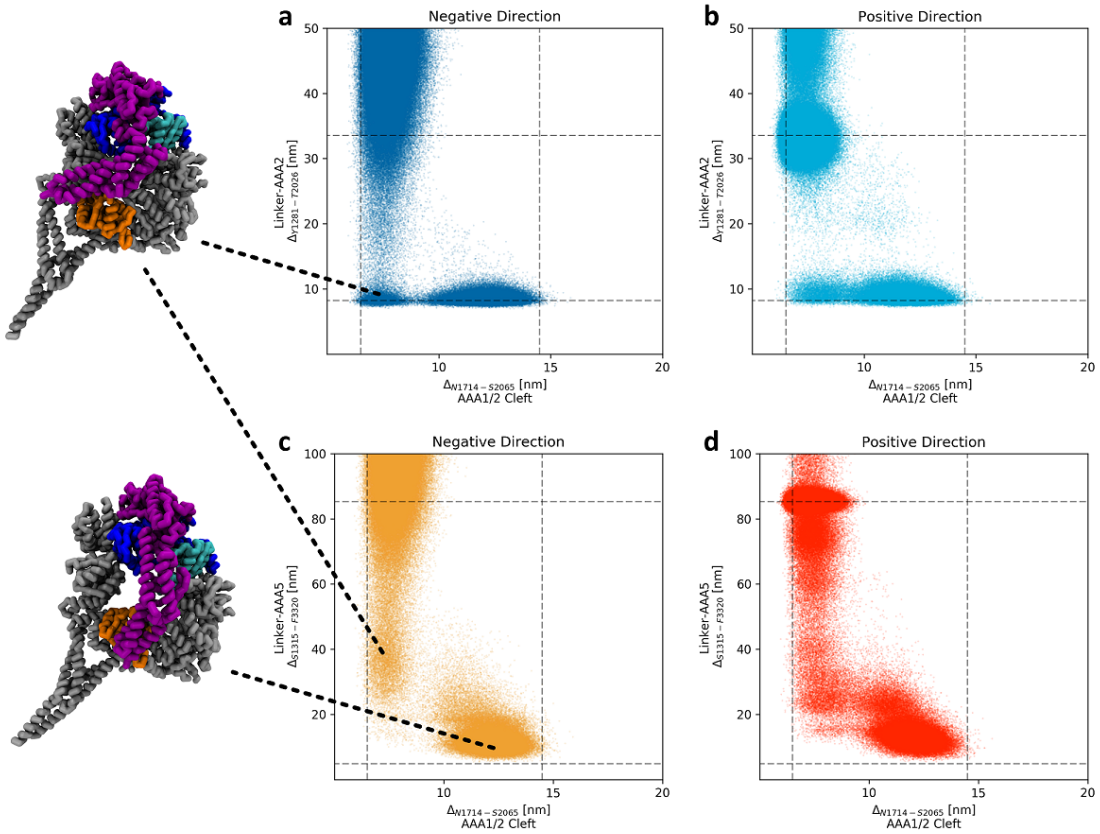


Figure 4.16: LN interactions under external load. a/b) Scatter plot of $\Delta_{Y1281-T2026}$ (LN-AAA2 distance) against $\Delta_{N1714-S2065}$ (AAA1/2 cleft) in 100 trajectories of dynein in the repressed state with increasing force applied in the negative (blue) and positive (cyan) directions. c/d) Scatter plot of $\Delta_{S1315-F3320}$ (LN-AAA5 distance) against $\Delta_{N1714-S2065}$ (AAA1/2 cleft) with increasing force applied in the negative (orange) and positive (red) directions. While the external force causes detachment of the LN from the AAA5 domain (c,d), the LN-AAA2 contact remains intact (a,b). The structure in the upper left shows the LN detached from the AAA5 domain but in contact with the AAA2 ILs. The structure in the lower left shows the LN bound to both AAA2 and AAA5 domains.

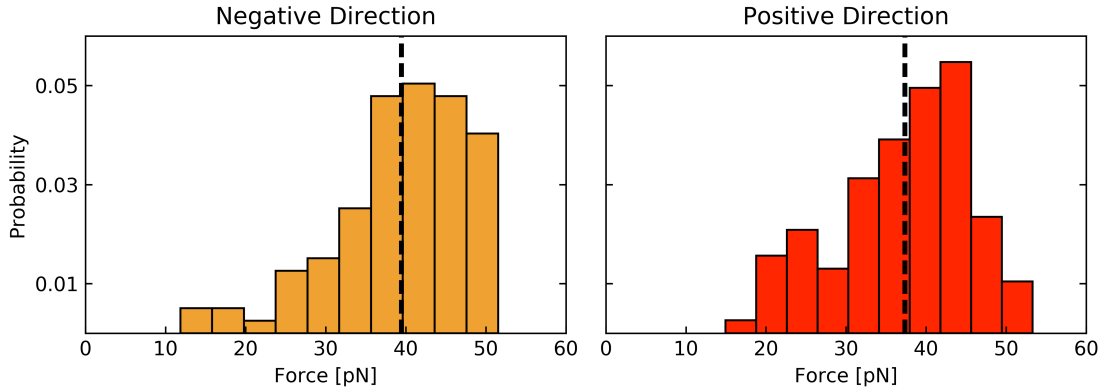


Figure 4.17: Mean unbinding force. Distribution of the force required for reactivation of the allosteric communication in the repressed state and detachment from MT. The distribution of force in the negative direction (orange) has a mean value of $39.4 \pm 8.2pN$. The distribution in the positive direction (red) has a mean value of $37.3 \pm 8.2pN$.

of external force on the motor. Interestingly, there was no significant difference between the two sets of simulations in terms of the magnitude of the unbinding force. Single molecule unbinding assays indicate that the mean unbinding force tends to depend on the directionality of the applied force[44,45,47]. However, this asymmetrical behavior presents itself also when the force is applied to the C-terminus of the motor, suggesting that the asymmetry is not LN dependent.

## 4.3 Discussion

The fact that the AAA2 ILs are imperative for dynein proper function is well established[17,18,43]. The precise mechanisms, however, with which the ILs regulate motor activity were not yet solved. Our simulations present a picture of the possible ways in which the AAA2 domain is involved in allosteric communication in the MD.

Of particular interest is the way in which ATP binding to the AAA1/2 pocket translates into detachment of the linker domain from its binding site on the AAA5 domain and the priming stroke transition. The effects of ATP binding on the

AAA1/2 domains can be separated into two parts. The first is a conformational transition of the ATP binding pocket, which occurs as a result of the interactions with the ATP molecule itself. These interactions are implicit in our simulations as we do not explicitly include the ATP molecule, but they are taken into account indirectly by the conformational information obtained from the crystal structure. The second part is the closure of the AAA1/2 cleft. This transition is likely due to the stabilization of the interactions between the AAA1 and AAA2 domains by the conformational change in the ATP binding pocket.

It is the closing motion of the AAA1/2 cleft that appears to transmit information to the rest of the molecule. In the particular context of the priming stroke, closing of the AAA1/2 cleft pulls the AAA2 domain towards the AAA1 domain and away from the LN. This leads to destabilization of the post stroke conformation as the AAA2 ILs are no longer in contact with the LN. The stabilizing interactions between the LN and the AAA5 domain are not strong enough for the two domains to stay in contact for long and the LN eventually undergoes a priming stroke.

The LN-AAA2 interactions also appear to play an important role in the repression of the allosteric pathway along the AAA1/2 and AAA5/6/S domains by the AAA3 unit. While it is clear that conformational changes in the AAA3 domain are involved in the repression of dynein activity, the manner in which they do so is no clear, in particular, because the changes are subtle[43]. Our simulations reveal that AAA3 unit represses the detachment from MT by over-stabilizing the interactions between the LN and the AAA2 ILs. In the crystal structure of the repressed state, the AAA3 and AAA4 domains are slightly shifted in relation to one another, when compared with other structures. This shift is enough to bring the AAA2 ILs closer to the LN. In addition to stabilizing the LN-AAA2 interactions, this shift may directly or indirectly destabilize the AAA1 and AAA2 interactions in the AAA1/2 cleft. As a result, the cleft is prevented from closing, even when ATP is bound

at the AAA1/2 binding pocket. As we have shown before, the LN and AAA2 ILs do come into contact regardless of the state of the AAA3 domain, however, these interactions by themselves are not enough to prevent conformational changes in the AAA1/2 domains.

Studies show that even when the AAA3 domain is occupied by ATP, if a strong enough force is applied at the N-terminus of the LN domain, dynein resumes normal activity[45]. The results of our simulations are consistent with this observations. Interestingly enough, pulling on the LN does not destabilize the interactions of the LN with the AAA2 domain. Instead, breaking the contacts between the LN and the AAA5 domain appear to be enough to reverse the effects of the AAA3 domain in the repressed state. This implies that the combination of LN interactions with both the AAA2 and AAA5 domains as well as the stabilization of these interactions by the ATP bound AAA3 domain are required for repression of dynein activity. It is sufficient to interfere with one of these three elements to allow dynein to undergo the priming stroke and detach from MT.

In the context of gating, it is worth pointing out that in our simulations, pulling in both the negative and positive directions along the MT axis had the same effect approximately. This implies that the AAA3 repression system does not work as a gating mechanism in the classical sense of creating an asymmetry between the leading and trailing MDs. It does work as a gating mechanism in that it prevents dynein detachment from MT in the absence of external strain, indicating that the repressed state may be important in the context of dynein processivity.

Since we know that dynein does respond asymmetrically to an external load, it might seem that our findings are incorrect. That is, however, not necessarily the case as there are other possible reasons for the asymmetric response. One such reason is the fact that when dynein binds to MT, it does so at an angle[21,71]. This, in and by itself, is enough to produce an asymmetric response to force, especially in

the context of single molecule unbinding experiments[44,45,47]. It has been pointed out that when the MDs in the dynein dimer are far enough, there is a higher probability that the trailing MD is going to step[13,14]. This indicates that there is some gating mechanism in dynein that is sensitive to internal strain. Our results do not account for this effect. A reasonable explanation, however, can be given by the possibility that a backward load, acting on the leading MD, would either prevent or slow down the leading motor's power stroke, which is necessary for ADP release and completion of the dynein mechanochemical cycle. This is a plausible gating mechanism that does not require an asymmetric response to force and is consistent with our findings.

# Chapter 5

# Kinesin: Overview

Members of the kinesin superfamily of motor proteins transport cargo along MTs towards their positive ends[6]. They are involved in a wide variety of processes that include the transportation of vesicles, mitochondria and neurite elongation components along axons as well as transportation of cargo within cells' bodies and within cilia and flagella[72,73,74]. Conventional kinesin, which is the focus of our research, and to which we refer to simply as kinesin, is involved in the transportation of cargo between the endoplasmic reticulum and the Golgi apparatus as well as the transportation of lysosomes and endosomes[75,76,77].

Kinesin is a dimeric protein. Each subunit contains a MD at its N-terminus followed by a neck linker (NL) and a coiled coil (CC) domain (see Fig. 5.1)[15,78]. The MD is the structural element of the motor responsible for interaction with the MT as well as the ATP hydrolysis cycle. Most of the motor's function is controlled through this domain. The NL connects the MD to the CC and is the structural component that performs the power stroke[9,79,80]. In addition it is thought to play an important role in the coordination between the MDs[81,82,83]. Finally, the CC, connecting the motor to the cargo, is the structural element to which the two MDs are joined.

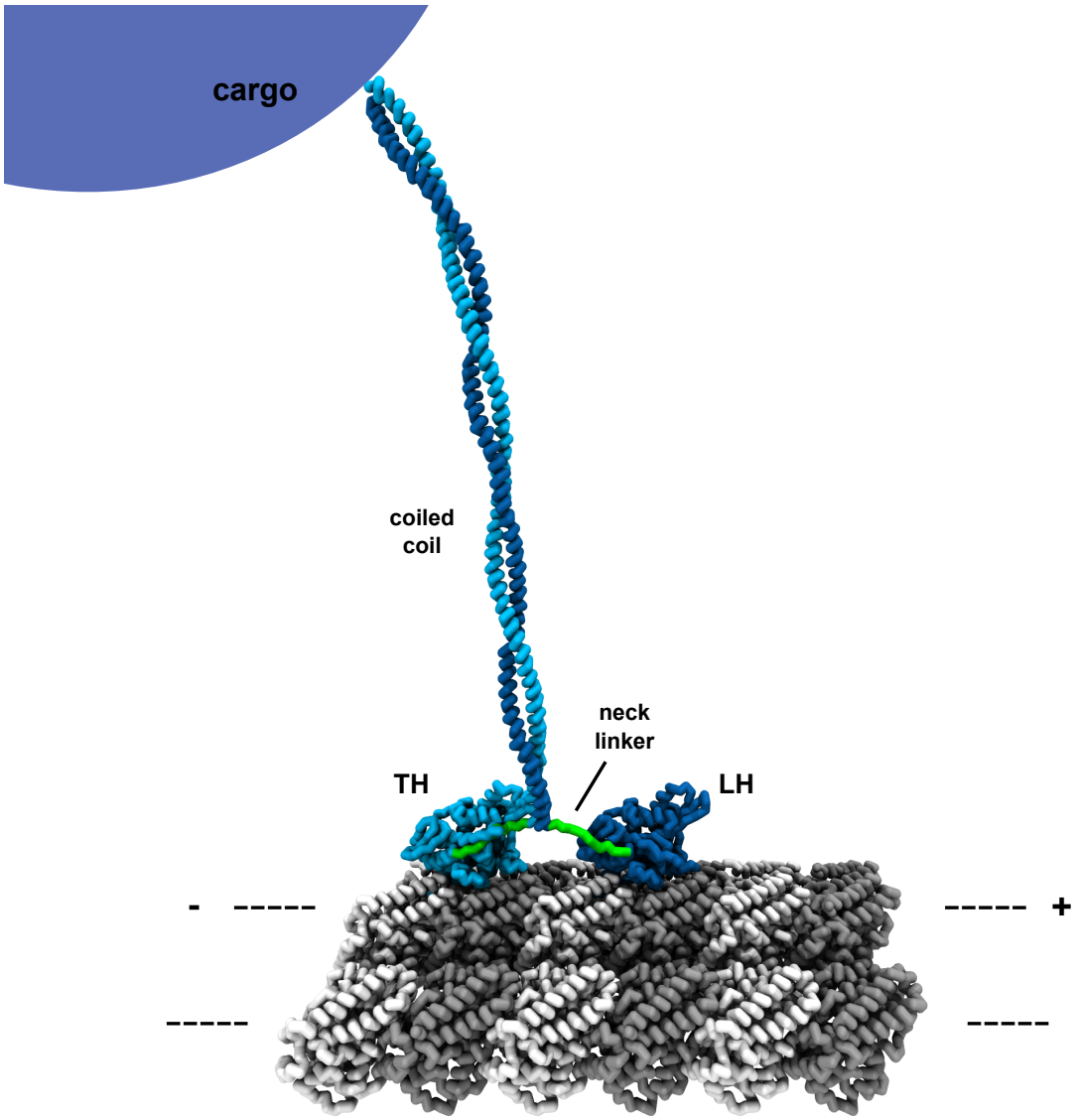


Figure 5.1: Structure of the kinesin-MT construct. The TH and LH are depicted in cyan and blue respectively. The NLs are colored in green. The $\alpha$ and $\beta$ tubulin units are colored in gray. The CC structure, extends from the NLs and binds to the cargo.

Kinesin moves along the MT, predominantly along a single protofilament, by performing alternating steps in which the MDs move in a hand over hand motion[63,64,84]. When resting between steps, both MDs are tightly bound to the MT. We refer to the MD that is closer to the MT's positive pole as the leading head (LH) and to the one closer to the negative pole as the trailing head (TH). The affinity of each MD for the MT is determined by its nucleotide state[58]. The mechanochemical cycle of kinesin is shown inf Fig. 5.2. The LH contains no nucleotide while the TH is bound to ADP and Pi. Upon release of the Pi, the MT affinity for the TH diminishes significantly and the TH detaches, thus initiating the step. Once detached, the TH diffuses in three dimensional space while being tethered to the bound LH by the two NLs. At this point the LH binds ATP, which induces a conformational change in the linker (the power stroke), docking it to the LH and biasing the motion of the detached MD towards the positive end of the MT[69]. Eventually the detached MD reaches the TBS along the protofilament. It has been shown that binding of the MD to MT accelerates the release of ADP from the MD[85]. Once ADP is released, the MD binds tightly to the TBS, becoming the new LH. Eventually the new TH hydrolyzes its ATP molecule, completing the cycle.

Kinesins step in a highly coordinated manner, indicating that the MDs are able to coordinate their cycles. Several studies over the last couple of decades indicate that strain between the two MDs affects the LH's affinity for the MT[81,82,83]. Furthermore, it appears that NL's orientation with respect to the MD determines the affinity[83]. Thus, the NL serves as the mechanical element with which the chemomechanical cycle of the LH is gated.

Much progress was made in recent years in our understanding of the mechanics of the motor's step itself. Both experimental and computational studies seem to indicate that the stepping MDs move by means of tethered diffusion towards the TBS[69,86]. The nature of the particular pathway taken by the stepping MD is still

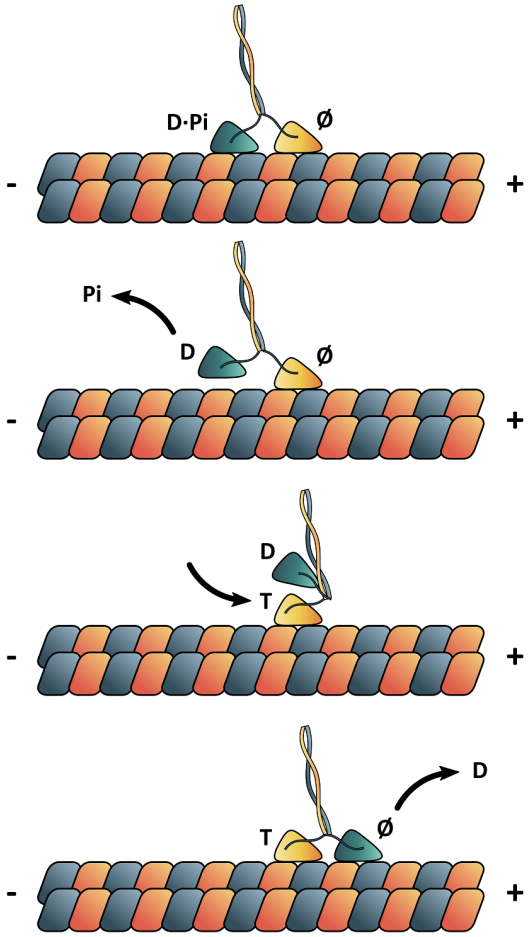


Figure 5.2: Chemomechanical cycle of kinesin. The cycle starts with the LH in the apo state and the TH is occupied by ADP and Pi. Upon Pi release, the TH detaches from the MT, allowing it to diffuse around the LH. Once ATP binds to the LH, its NL undergoes a power stroke by docking to the MD. This biases the motion of the TH towards the positive end of the MT. Once the TH releases its ADP molecule, its affinity for MT increases and it rapidly binds to its TBS, making it the new LH and completing the cycle.

somewhat under debate. There is evidence that suggests that the motor steps in an asymmetric hand over hand fashion, that is, it alternates in the way the stepping MD moves around the LH, clockwise and counterclockwise. At the same time, there are studies that indicate that the motion is of the symmetric kind and that the stepping MD pauses predominantly on one side of the LH. In this work we explain the discrepancy between these two observations.

# Chapter 6

# Effects of Cargo on the Stepping Mechanism of Kinesin

## 6.1 Introduction

The hand-over-hand model for the walking mechanism of kinesin is well supported by experiment[63,84]. According to the model, the TH detaches from the MT, diffuses over the non stepping LH, and eventually binds to the MT $\approx 16nm$ toward the positive end, thus becoming the new LH. What is less clear is what are the specific pathways taken by the stepping head. More specifically, because of the geometry of the motor structure (see Fig. 5.1), the TH has to rotate around the CC in order to pass the LH. This leads to two possible scenarios. In the first scenario the MD passes the LH on its right, when looking toward the MT's positive end (Fig. 6.1). In the second scenario the TH passes the LH on the left.

There is debate in the motor community as to whether the motor steps repeatedly on the same side or whether it alternates between the two sides. The model in which the motor steps along the same side each time is referred to as the symmetric hand-over-hand model[84]. This is due to the fact that each step is geometrically

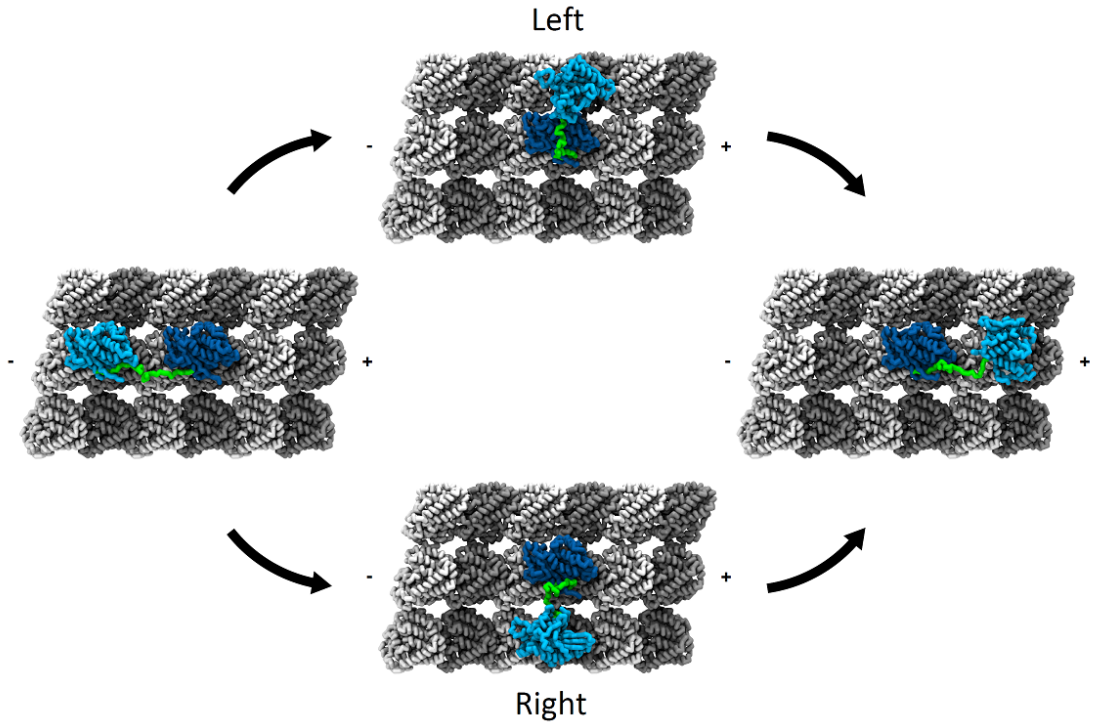


Figure 6.1: The kinesin TH can complete a step through one of two pathways. It can either step through the left side (top pathway) or step through the right side (bottom pathway). According to the symmetric hand-over-hand model, the TH would step trough the right side on every step. In contrast, in the asymmetric model the motor would alternate between right and left on each step.

identical. Similarly, the model in which the motor alternates between right and left steps is referred to as the asymmetric hand-over-hand model. Whether kinesin steps in a symmetric or asymmetric manner has implications regarding tension along the CC. If kinesin steps in a symmetric manner, the CC would be rotated by 180 deg per step, leading to accumulation of rotational stress along the CC and to potential rotation of the kinesin's cargo.

Several studies indicate that kinesin steps in an asymmetric fashion. By tracking the rotational motion of a cargo bound to kinesin, it has been shown that there is no cumulative rotation, suggesting that the motor does not generate rotational tension[87]. Furthermore, Asbury *et al* show that in two consecutive steps, the mean dwell times are different, yet again indicating that kinesin steps asymmetrically[84]. Nevertheless, recent studies in which the position of the TH was tracked with relatively high precision, indicate that the stepping motor is predominantly found on the right side of the LH[88]. This suggest that kinesin steps symmetrically.

In order to better understand the details of the kinesin stepping mechanism and to reconcile the two seemingly contradictory scenarios, we conducted CG simulations of the kinesin step which have proven useful in the description of molecular motors[68,69,89]. We show that the stepping pattern is highly dependent on the experimental setup. In the presence of the cargo there is a preference towards the asymmetric model. However, when the orientation of the CC is not constrained, we show that there is a high level of stochasticity in the directionality of the stepping motion.

## 6.2 Results

### 6.2.1 Stepping Pattern of Kinesin with Free Rotation of the CC

We first set out to find the stepping pattern of the kinesin motor when the CC is attached to a cargo but no rotational constraint is imposed on it (Fig. 6.2). We generated 100 independent trajectories of the kinesin stepping process for $18\mu s$ using our CG model, which has been shown to successfully capture the kinematics of the kinesin step[69]. During the simulations we monitored the position of the center of mass of the TH as a function of time. In their 2016 study, Isojima *et al* have been able to track a single kinesin motor head and generate trajectories on a two dimensional plane[88]. In order to compare our simulations with their results, we defined a two dimensional plane, $\mathcal{P}$ (See appendix C), as the plane perpendicular to the vector $\mathbf{n}$ which points from the center of the MT towards the protofilament on which the kinesin motor walks. We then projected the position of the center of mass of the TH onto $\mathcal{P}$.

The results, shown in Fig. 6.3, agree with recent experiments, indicating that the TH is found predominantly on the right side of the LH during the stepping process[88]. The probability of the TH to be on right side of the LH as a function of time shows that after about $5\mu s$, in approximately 90% of trajectories, the TH is on the right side of the LH (Fig. 6.4).

Our findings seem to support the symmetric hand-over-hand model for the kinesin stepping pattern, which appear to contradict the findings of two pioneering experimental studies[87,90]. In order to further investigate whether our results support the symmetric model we measured the end-to-end vector of the TH linker,$\mathbf{r}_{NL}$ , projected onto $\mathcal{P}$, as a function of time (Fig. 6.5). This vector indicates the relative

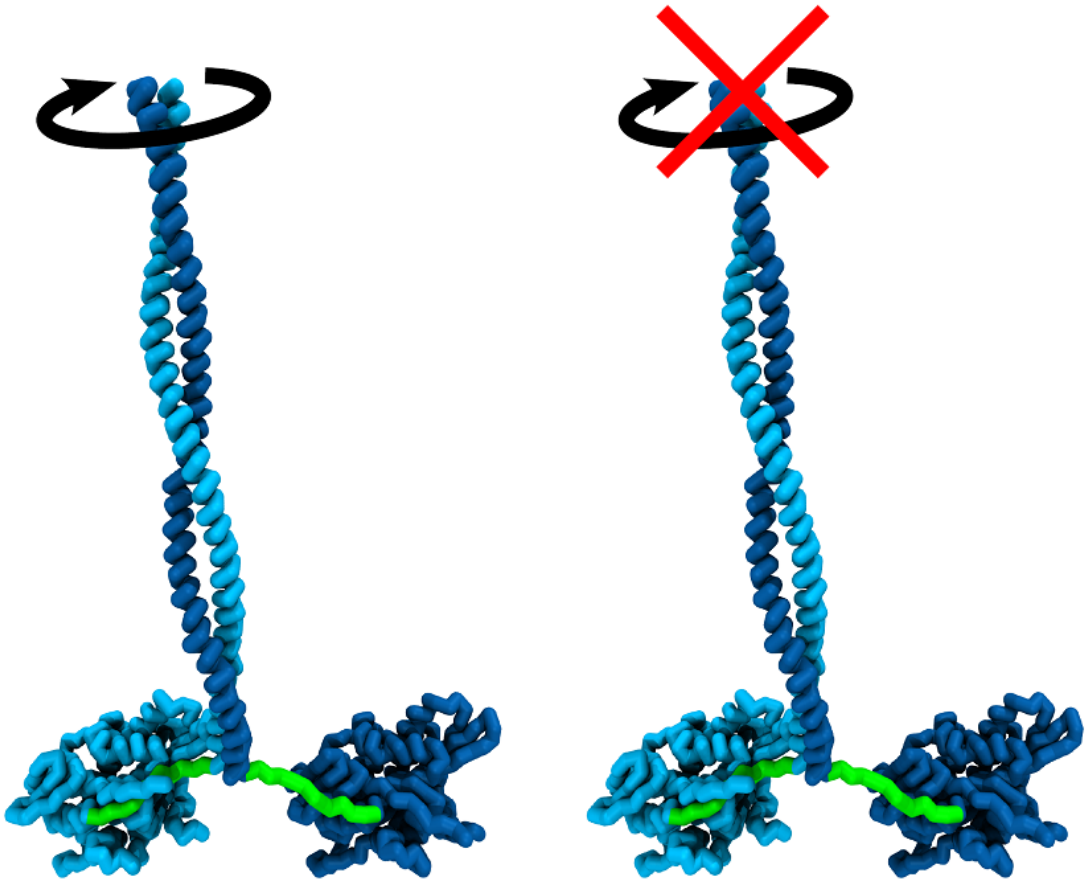

Figure 6.2: The left structure depicts a kinesin motor in which the CC is allowed to rotate freely despite being bound to a cargo. In contrast, the right structure depicts a kinesin motor in which the top of the CC is constrained in such a manner so that its orientation is fixed in time.

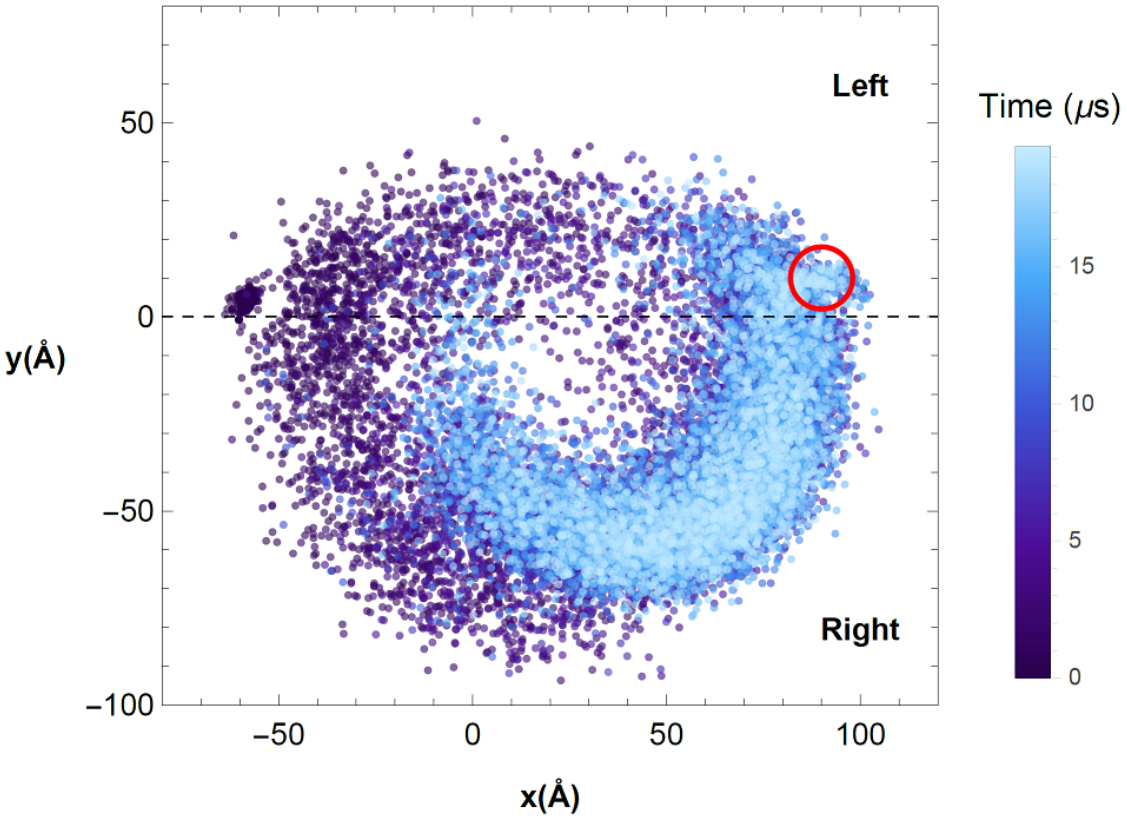


Figure 6.3: Projection of the position of the center of mass of the TH onto the plane $\mathcal{P}$, which is explained in the SI. The data points correspond to the ensemble of trajectories over time where the CC, despite being attached to the cargo, was rotationally unconstrained (left structure in Fig. 6.2). The 0 value along the y axis (the position of the center of mass of the LH) divides the system into left (positive values along the y axis) and right (negative values along the y axis) sides. The color of each data point corresponds to the time in which the center of mass of the TH is in a particular configuration. The scale for time can be found on the right side of the figure. The red circle marks the positions of the TBS.

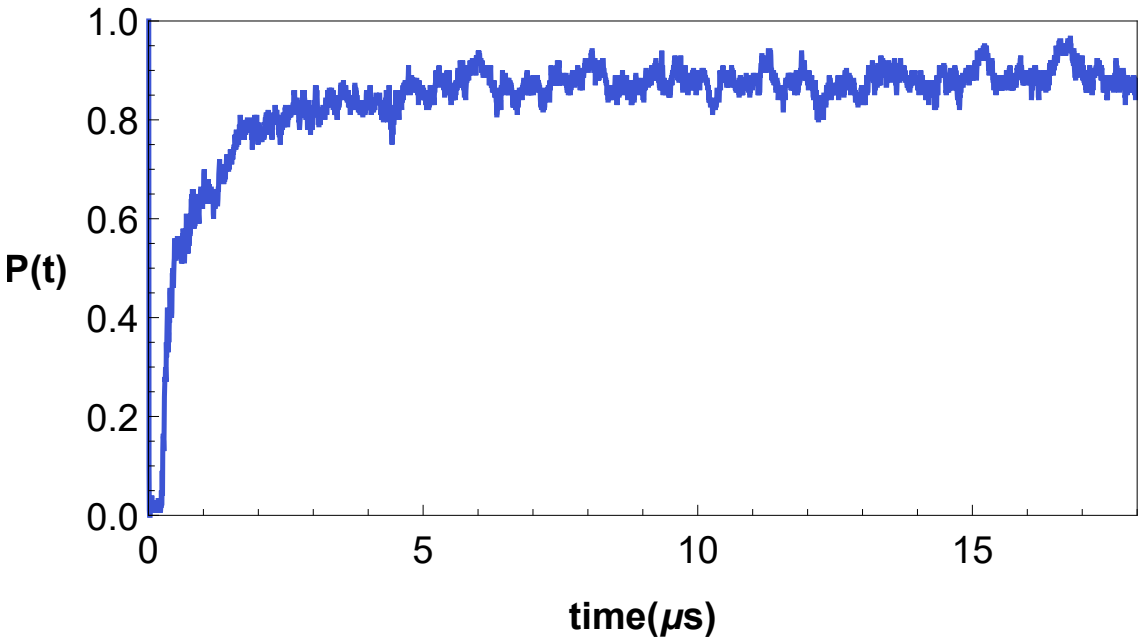


Figure 6.4: The probability, $P(t)$, that the TH to be found on the right side of the LH as a function of time when the CC is not rotationally constrained.

orientation between the two motor heads during the stepping process. We then calculated the angle $\theta$, which is the angle between $\mathbf{r}_{NL}(t)$ and $\mathbf{r}_{NL}(0)$, in order to determine whether the THs tends to pass the LH on either the right or left side. The results show that, while the TH is found most of the time on the right side of the LH, in a significant portion of the trajectories, the TH reaches that position by traversing the LH from its left side (Fig. 6.6). In fact, only in 69% of the trajectories does the TH step from the right side. This percentage is also reflected in the number of trajectories in which the TH reaches the TBS. Out of 13 trajectories in which the step is completed, in 4 of them the TH does so by passing the LH on the left, even though there is no rotational constraint on the CC.

This surprising result has two main implications. The first is that the position of the TH relative to the LH is not a good indicator of the pathway taken by the TH. One reason for this is that the left pathway is located above the LH rather than to its left. Furthermore, while the likelihood of the TH traveling through the

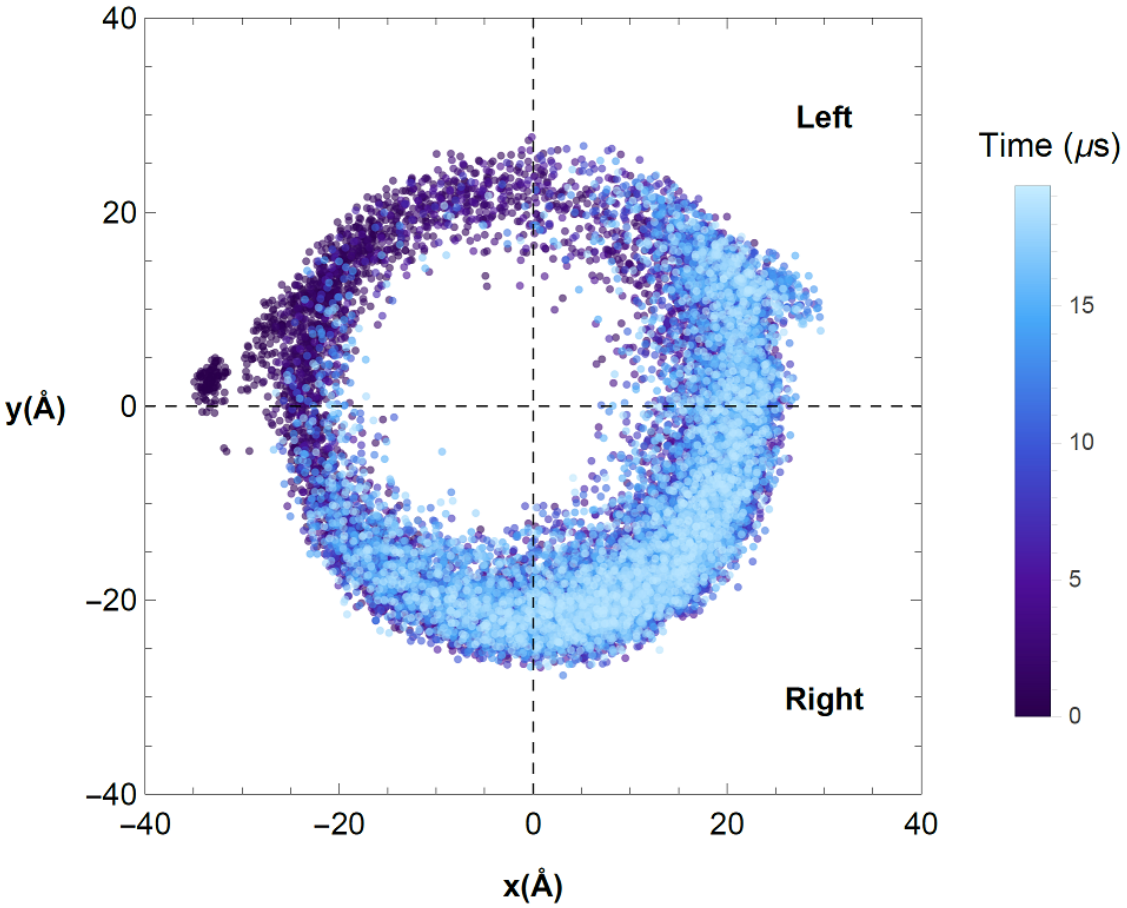


Figure 6.5: Projection of $\mathbf{r}_{NL}$ onto $\mathcal{P}$ for the unconstrained ensemble. The data points presented have been time smoothed to allow for easier calculation of $\theta$.

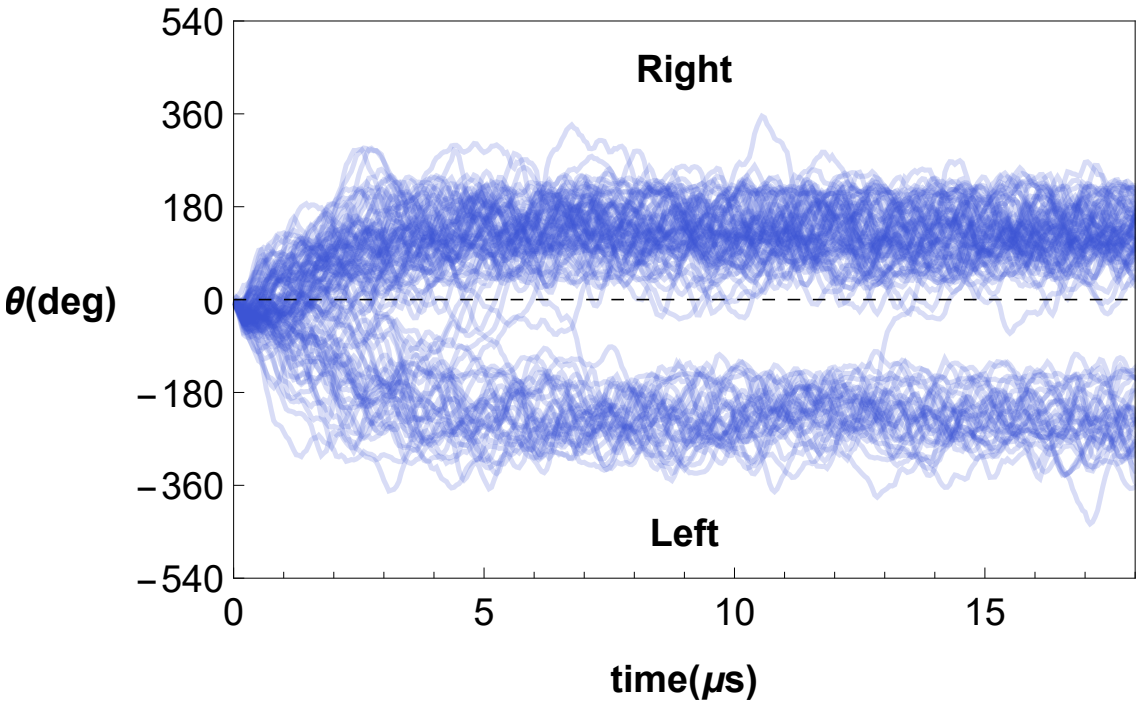


Figure 6.6: Values of $\theta$ in the trajectory ensemble over time where $\theta$ is the orientation between the two motor heads (see appendix C). In these simulations the CC was not rotationally constrained.

left pathway is not small, the probability of finding the TH on the left side of the leading is, as can be seen from Fig. 6.3. This implies that the transitions through the left pathway are short lived, making them hard to detect.

The second implication is that when no rotational constraint is applied to the CC, the kinesin motor can step in both a symmetric and an asymmetric fashion. Given the probabilities obtained from our simulations, and assuming that the steps are uncorrelated, the probability of the motor to step twice from the same side consecutively is approximately 57% (see appendix C), suggesting a slight preference for the symmetric model. As a result we would expect the kinesin motor to exert only a small torque on the CC and through it on the cargo during the run along the MT. Nevertheless, we do expect the motor to exert some amount of torque on the CC between individual steps since in each step there is a 180° rotation of the motor.

### 6.2.2 Cargo Rotation is Dominated by Thermal Fluctuations

Single molecule studies of kinesin have shown, by tracking the rotational motion of a cargo attached to the motor, that kinesin does not step in a symmetrical hand-over-hand fashion[87,90]. If it did so then there should be, on an average, a 180° rotation of the cargo in the same direction per step. They observed that the cargo does not exhibit such a cumulative rotation, as expected from such a scenario, and therefore concluded that the stepping mechanism is asymmetric.

Even though our simulations show that the results obtained by Tomishige and co-workers do not necessarily imply a symmetric stepping mechanism, we still find that kinesin can step in a symmetric manner[88]. In order to reconcile our results with experimental observation we pose an inverse question. Instead of asking how would the stepping mechanism manifest itself in the rotational motion of the cargo, we explore the effects of the cargo on the stepping mechanism. More specifically, we ask whether the motor determines the rotation of the cargo or does the cargo determine the stepping pattern of the motor? To answer this question, we need to assess how much torque should the motor have to exert on the cargo in order to produce significant rotation that is experimentally detectable.

Gutierrez-Medina *et al* obtained an estimate of the rotational diffusion constant of the $1.3\mu m$ cargo bead used in their single molecule experimental set up and found it to be $D_r = 6.8 \times 10^{-2} rad^2 \cdot s^{-1}$ or roughly $D_r \approx 223 deg^2 \cdot s^{-1}$[87]. Assuming that the kinesin CC is relatively rotationally stiff, the torque that the motor exerts on the cargo can be approximated by the product $N_{mot} = r_{CC} \times F_{ten}$ where $r_{CC}$ is the radius of the CC structure and $F_{ten}$ is the tension applied to the base of the CC by the linkers of the two MDs bound to the MT.

The value of $r_{CC}$, which is easily obtained from crystal structures, is about

0.5nm. An accurate value for the inter head tension, $F_{ten}$, is not available. However, we can put an upper bound on its value using the following argument. Uemura *et al* have shown that the mean unbinding force of a single kinesin head that is bound to a MT with high affinity ranges between 6-10 pN approximately[58]. Since the two head bound state is relatively stable in a nucleotide free environment, it is unlikely that the tension between the two heads exceeds 10 pN since that would lead to rapid dissociation of one of the motor heads from the MT. Other estimates of the tension fall in the 10-15 pN range which is not much larger in comparison[91]. We use $F_{ten} = 10pN$ as a reasonable upper bound.

Given these estimates, we conclude that the angular drift velocity is estimated as, $\omega_d = \frac{N_{mot}}{\zeta}$, where $\zeta$ is the friction coefficient associated with the cargo, which can be obtained using the relation, $\zeta = \frac{k_B T}{D_r}$. Our estimate of the angular drift velocity is $\omega_d \approx 4.7 deg \cdot s^{-1}$. In single molecule experiments that are conducted at low ATP concentration, the order of magnitude of the dwell time between steps is typically seconds. The standard deviation of the angular fluctuations due to thermal noise at a one second time scale can be estimated as, $\sigma_\theta = \sqrt{2D_r * 1s} = 21.1 deg$, which is 4-5 times larger than the calculated contribution of the drift term during the same time period. Given the ratio between the thermal and drift contribution to the angular motion of the cargo, even though the drift contribution cannot be neglected, we can assert that the rotational motion of the cargo is dominated by the thermal fluctuations during the waiting time between steps. Furthermore, the torque that is generated by the motor can only produce a small cargo rotation of a few degrees, which is almost negligible in comparison to the full 180° rotation that is needed to dissipate the rotational strain along the CC. This implies that the rotational strain has to be resolved at the motor level and not by rotation of the cargo. Given these facts, we set out to answer the following question. How does the torque along the CC affect the stepping pattern of the kinesin motor?

It is important to note that in the case of the experiments that were carried out by Gelles and co-workers, the argument that the torque generated by the motor is too small to produce a large rotation in the cargo does not hold[90]. This is because our conclusion is based on the assumption that the cargo is approximately spherical. In the above mentioned study, however, measurement of CC rotation was done by tracking the rotation of a MT attached to an immobilized kinesin. We expect a MT of the dimensions used in the experiment to have a smaller rotational friction coefficient in comparison to a bead. Furthermore, the experiment was carried out at very low concentrations of ATP, which ensured that the waiting time between steps was on the order of tens of seconds. Such long waiting times are significantly larger than the typical relaxation time for the rotational diffusion of the MT. The important conclusion was that kinesin does not walk in a symmetric hand-over-hand motion since no cumulative rotation of the MT was observed. In fact, rotations of 180° were rarely observed if at all.

This result is surprising since it has been well established that kinesin steps in a hand-over-hand mechanism, irrespective of whether the motion is symmetric or asymmetric[63]. A hand-over-hand mechanism implies that we would expect to observe a 180° rotation of the CC per turn regardless of directionality. The difference between the symmetric and the asymmetric models is that in the symmetric model each turn adds to the total rotation of the CC where as in the asymmetric model each two consecutive turn cancel each other. The fact that Hua *et al* do not detect such turns at all leads to a contradiction with the hand-over-hand model. A possible explanation for these results can be given by noting that, at very low ATP concentrations, kinesin has been shown to be found predominantly in the one head bound state while occasionally transitioning back and forth to a two head bound state[92]. Such transitions would allow for full dissipation of the rotational strain along the CC at the motor level without generating cargo rotation. Since this behavior is

highly non-trivial, we focus on systems with spherical cargo only.

### 6.2.3 Rotational Constraints of the CC Affect Stepping Mechanics

Given the observation that the torque generated by the motor does not manifest itself in rotational motion of the cargo beyond what is estimated due to thermal fluctuations, we expect the torque buildup resulting from one step to affect the next step of the motor. These conditions favor the asymmetric model since each step would cancel the torque generated by the previous step.

In order to quantify these effects we simulated two consecutive steps of the kinesin motor while fixing the orientation of the cargo end of the CC (Fig. 6.2). This constraint mimics the effect of the cargo not changing orientation on an average between steps. In order to generate our starting configuration we first equilibrated the system with no constraint on the CC, and generated an ensemble of orientations. We then switched on the constraint and generated 99 stepping trajectories. Out of these trajectories, we chose a single trajectory in which the motor completed a step by traversing the LH on the left side and used the final system configuration in this trajectory as our starting point for the next step.

Given these new starting conditions we generated 100 new trajectories, which we refer to as the first step, for $18\mu s$. We tracked the center of mass of the TH, projected onto $\mathcal{P}$, as a function of time (Fig. 6.7). When comparing the results with those for kinesin with no constraints on the CC (Fig. 6.3), we observe that in the constrained case the left side is less populated at early times and the right side is more populated in general. This is supported by our measurement of the probability of the motor being on the right side as a function of time (Fig. 6.8).

We calculated $\theta$ as a function of time for all trajectories In order to monitor on which side does the TH step (Fig. 6.8). In contrast to the unconstrained motor,

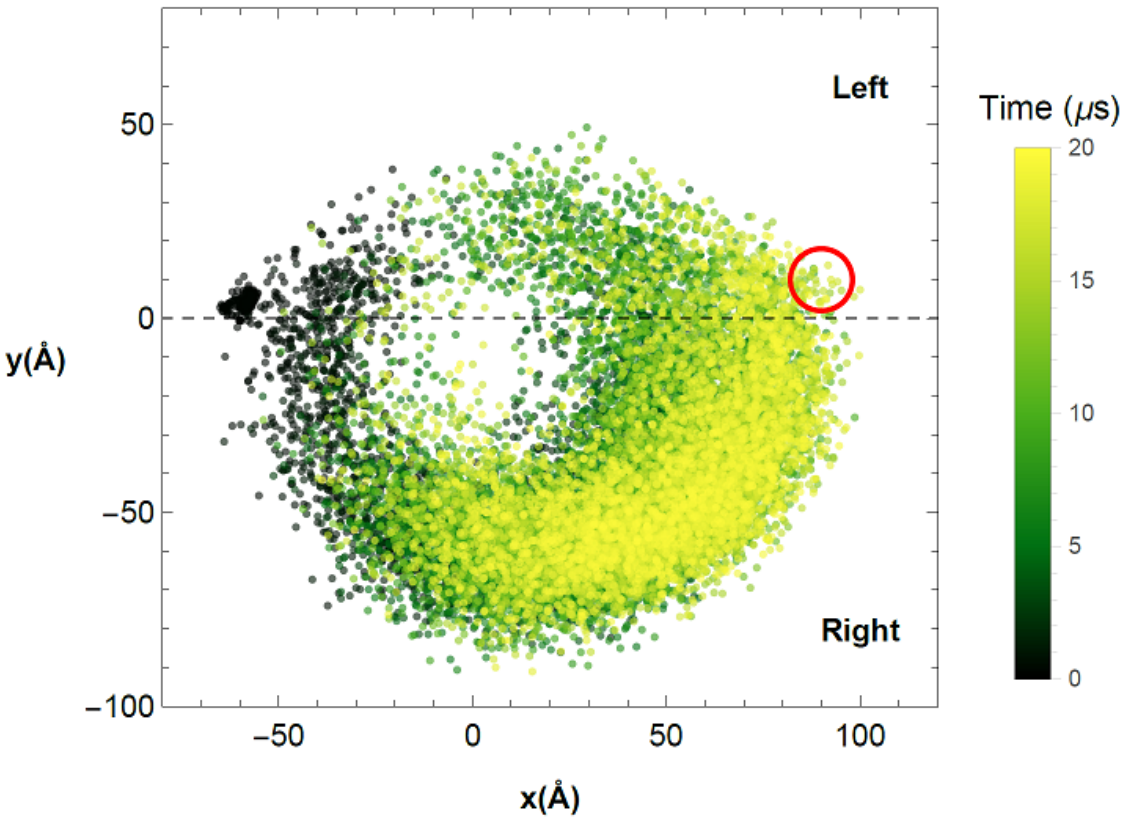


Figure 6.7: Projection of the position of the center of mass of the TH onto the plane $\mathcal{P}$ during the first step. The data points correspond to the ensemble of trajectories over time where the CC is rotationally constrained. The color of each data point corresponds to the time in which the center of mass of the TH is in a particular configuration. The scale for time can be found on the right side of the figure. The red circle marks the positions of the TBS.

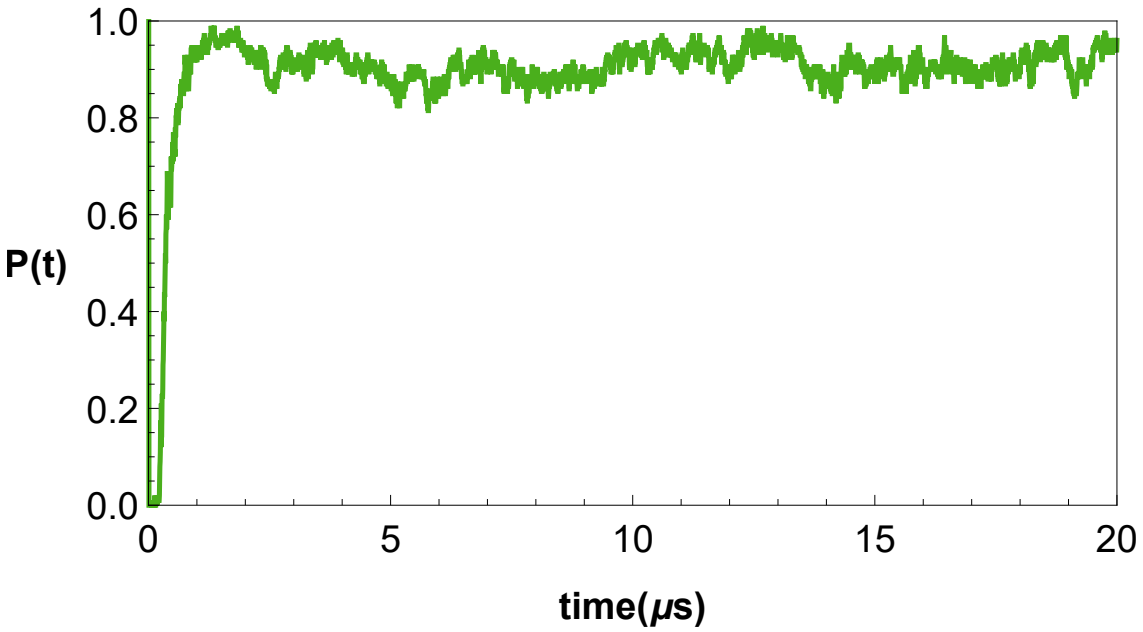


Figure 6.8: The probability of the TH to be found on the right side of the LH as a function of time during the first step of the rotationally constrained motor.

essentially in all the trajectories the TH passed the LH from its right. This was the case also for all the trajectories in which the TH reached its TBS because, the starting configuration of the system was such that the current step followed a step from the left. We would expect a bias towards the right side due to the toque buildup in the CC.

In order to investigate further we simulated an additional step which we refer to as the second step. This time we used a completed step from one of the first step trajectories as the initial configuration. We generated an additional 99 trajectories for $48\mu s$ and monitored the position of the TH's center of mass, projected onto $\mathcal{P}$, as a function of time (Fig. 6.10). The results were dramatically different from the previous step. The most striking difference was the tendency of the TH to be found closer to the original binding site at the negative pole of the MT. This is despite the fact that the docked NL should bias its motion towards the plus end of the MT. The docking of the NL does shift the overall position of the motor towards the + end. However, the orientation of the TH's NL, $\mathbf{r}_{NL}$, is biased towards the - end due

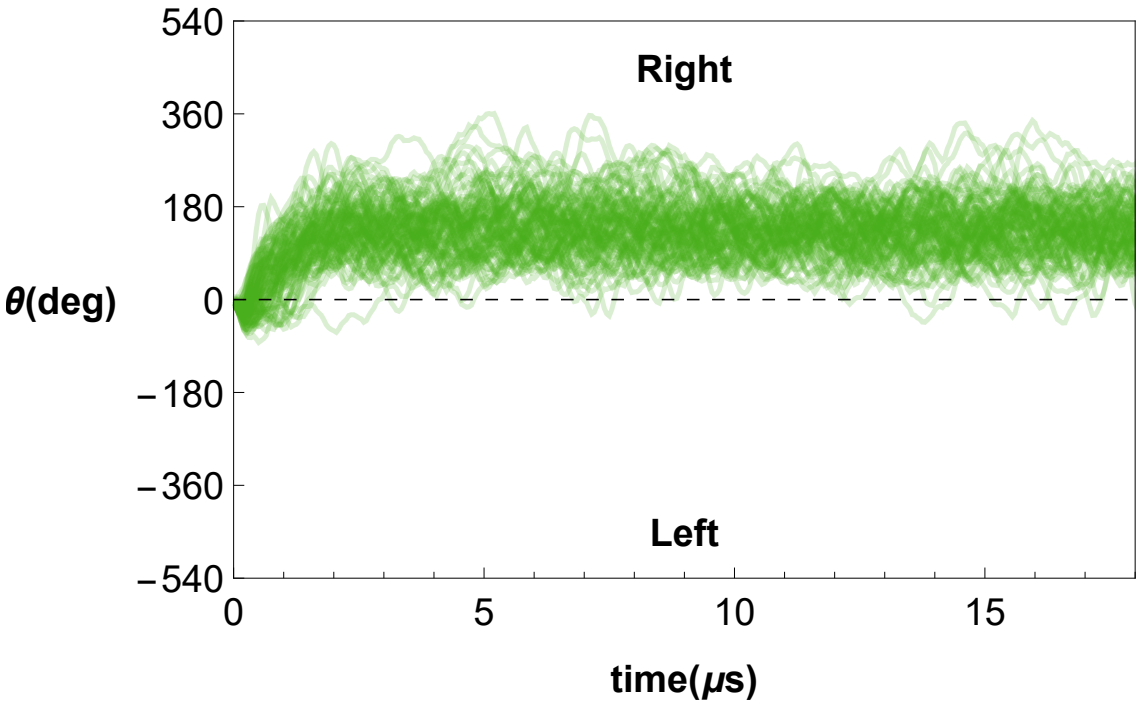


Figure 6.9: Ensemble values of $\theta$ over time during the first step of the rotationally constrained motor. $\theta$ represents the orientation between the two motor heads.

to the torque buildup (Fig. 6.11).

This effect also manifests itself in the duration of the stepping process. While in our previous simulations, in several trajectories the TH reached the TBS within $18\mu s$, in the current step we observed only one such event. Furthermore, the TH did not reach the TBS until after more than $40\mu s$.

We monitored the probability of the TH to be found on the right side as a function of time (Fig. 6.12). While there was still a clear preference for the right side, the probability was noticeably smaller in comparison to our previous simulations, particularly at early times. We proceeded to calculate $\theta$ as a function of time for the ensemble (Fig. 6.13). Interestingly, the simulations showed a small bias towards the right side over time, although the distributions seemed to be more centered around zero. Furthermore, there was no split between two clear populations as was the case for the unconstrained simulations.

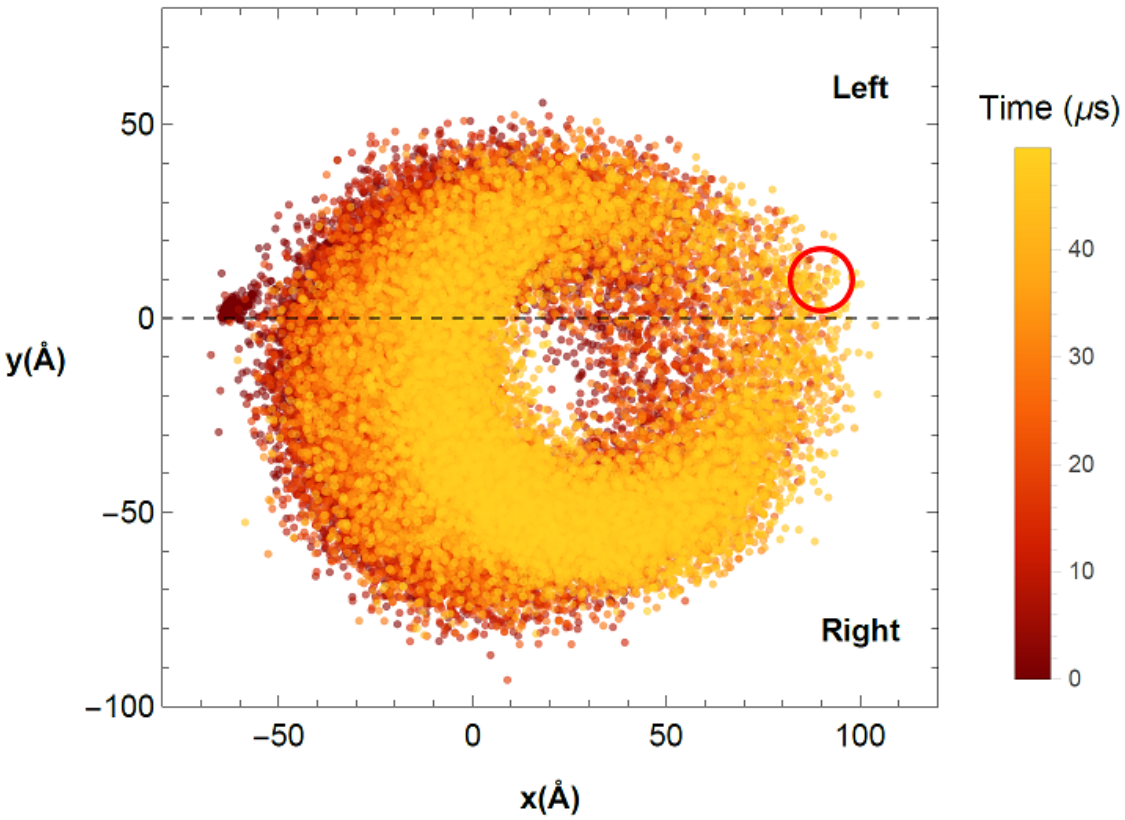


Figure 6.10: Projection of the position of the center of mass of the TH onto the plane $\mathcal{P}$ during the second step. The data points correspond to the ensemble of trajectories over time where the CC is rotationally constrained. The color of each data point corresponds to the time in which the center of mass of the TH is in a particular configuration. The scale for time can be found on the right side of the figure. The red circle marks the positions of the TBS.

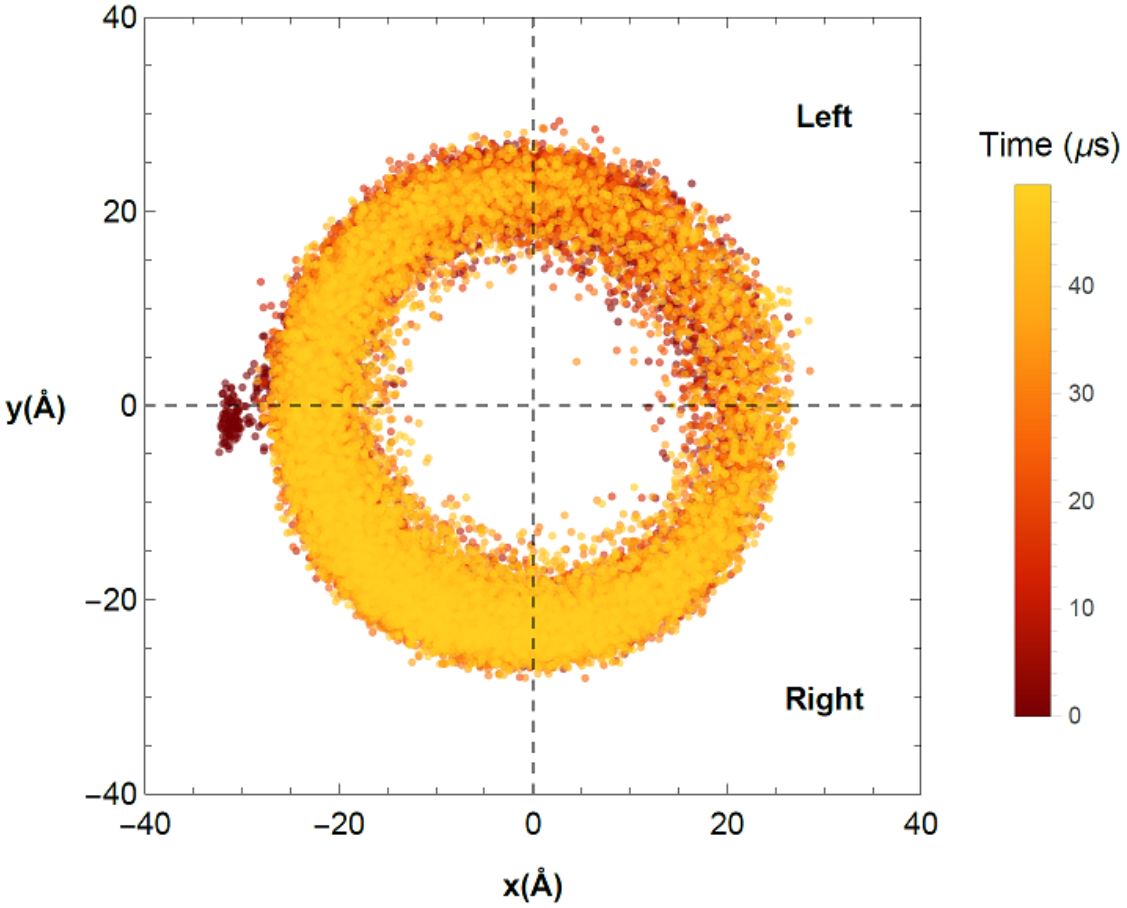


Figure 6.11: Projection of $\mathbf{r}_{NL}$ onto $\mathcal{P}$ for the constrained ensemble during the second step. The data points presented have been time smoothed to allow for easier calculation of $\theta$.

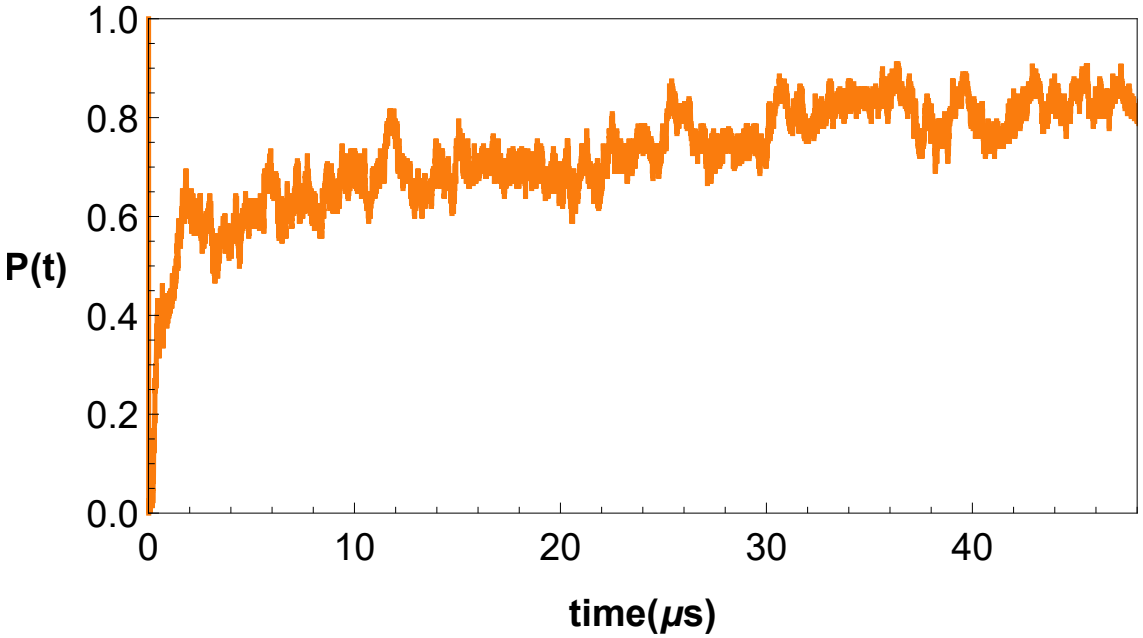


Figure 6.12: The probability of the TH to be found on the right side of the LH as a function of time during the second step of the rotationally constrained motor.

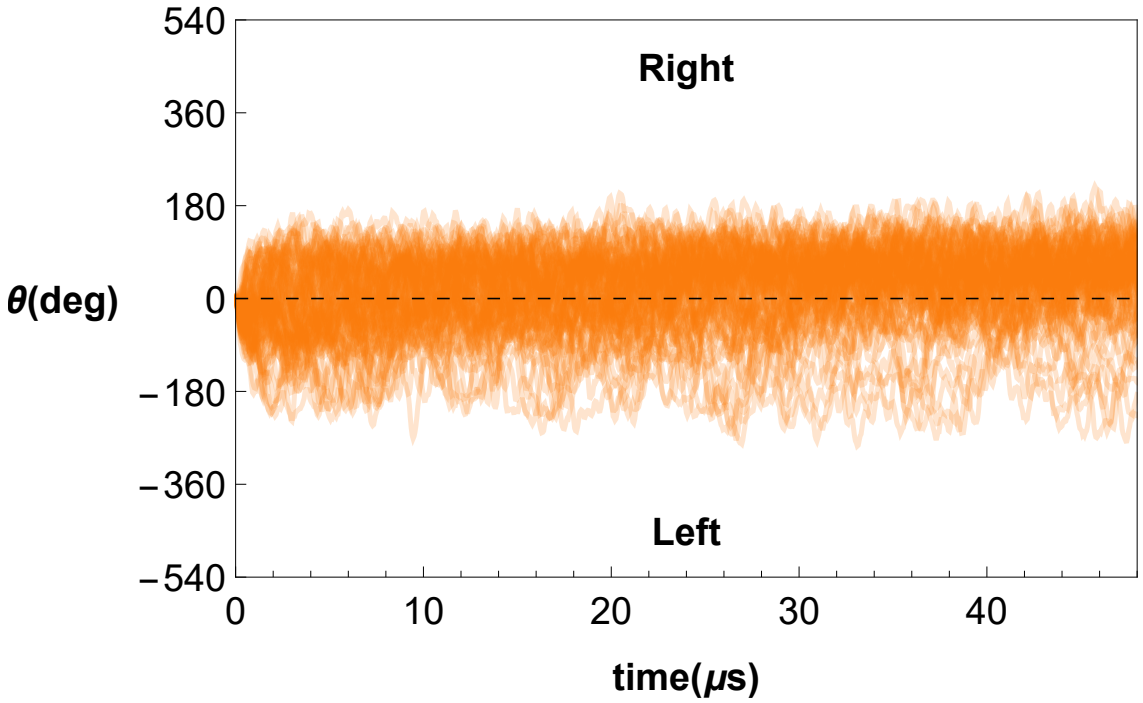


Figure 6.13: Ensemble values of $\theta$ over time during the second step of the rotationally constrained motor. $\theta$ represents the orientation between the two motor heads.

These results are somewhat surprising since one would expect that, according to the asymmetric model, following a step from the right, there would be a clear preference for stepping on the left. We did observe an overall rotation of the orientation distribution (Fig. 6.11) but not enough to clearly bias the system towards steps from the left.

In order to investigate further we computed the probability histograms of $\theta$ at the last frame of the trajectory for all three simulations sets (Fig. 6.14). Based on the non constrained ensemble measurements of $\mathbf{r}_{NL}$ (Fig. 6.2), the typical values of $\theta$ for the THs that reached the TBS are around $-150°/210°$ depending on whether the motor steps from the left or right respectively. In the unconstrained simulations these values are clearly reached (Fig. 6.14a). Same is true for the right side value in the simulations of the first constrained step (Fig. 6.14b). Interestingly, when looking at the $\theta$ distribution of the second constrained step (Fig. 6.14c), there is a clear preference for the TH to travel on the right side. However, given that in a particular trajectory the TH traveled on the left side, it was more likely to reach the neighborhood of $\theta$ values that allows binding to the TBS in comparison with those traveling on the right side. It is also important to note that in the one trajectory in which the TH reached the TBS, it did so by traveling on the left side.

Based on these findings, it is possible that the probability of reaching the TBS by traveling through the left side is higher even though the event of the TH having a positive $\theta$ value is more likely. This scenario is consistent with the asymmetric model.

### 6.2.4 CC Orientation Affects NL Tension

One of the main arguments in support of the asymmetric model was made in a study by Block and co-workers[84]. In their work, Asbury *et al* describe how the dwell times of kinesin can alternate between two distinguishable time scale from one step to the

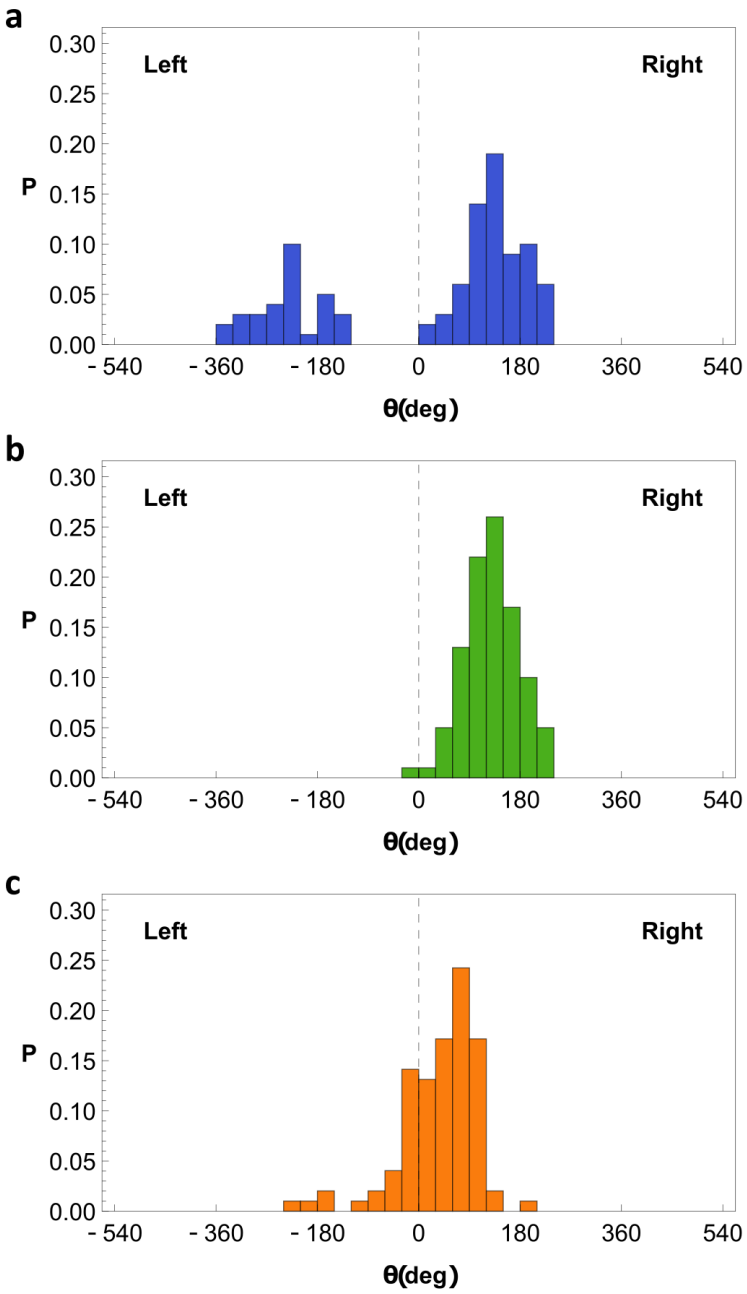


Figure 6.14: Histogram of the values of $\theta$ at the end of the simulations. a) The histogram in blue corresponds to the simulations of the unconstrained motor. b) The histogram in green corresponds to the simulations of the first step with rotational constraints. c) The histogram in orange corresponds to the simulations of the second step with rotational constraints.

next. They described this behavior as a sort of "limping" and took it as evidence that the motor alternates between two stepping conformations.

A possible cause of these changes in the dwell time is a change in the tension along the NLs due to the torque applied by the CC. Since a difference in the force that is applied to each motor head would affect its detachment rate, it is reasonable to assume that this is the cause of the "limping" behavior. This is further supported by the fact that this "limping" behavior diminished as the CC was extended[84]. In the case of a long CC, the CC may be able to absorb some of the torque.

In order to assess whether this is a plausible theory, we computed the distribution of tension along the NLs during the waiting times between the steps in our simulations. We made three separate calculations. One for the initial conformation of the unconstrained system and two using the starting configurations of the first and second constrained steps.

As can be seen from Fig. 6.15, the tension in the unconstrained system was lower in comparison with the constrained systems. This is expected since there is no constrain acting on the CC and therefore no torque is applied which can manifest itself in an increase in tension. Interestingly, when comparing the tension distributions of the first and second constrained steps (Fig. 6.15b and 6.15c), there is a clear difference in the mean tension. This is consistent with the picture presented by Asbury *et al*[84].

It is important to note that the forces obtained through our simulations are large in comparison to the range of forces we would expect, which are on the order of 10 pN. This is likely because of the use of a simple model for kinesin stepping. Nevertheless, the qualitative behavior of the forces found in our simulations is consistent with the experimental observations made by Block and co-workers, which in turn makes us confident that our model captures the behavior of the system correctly.

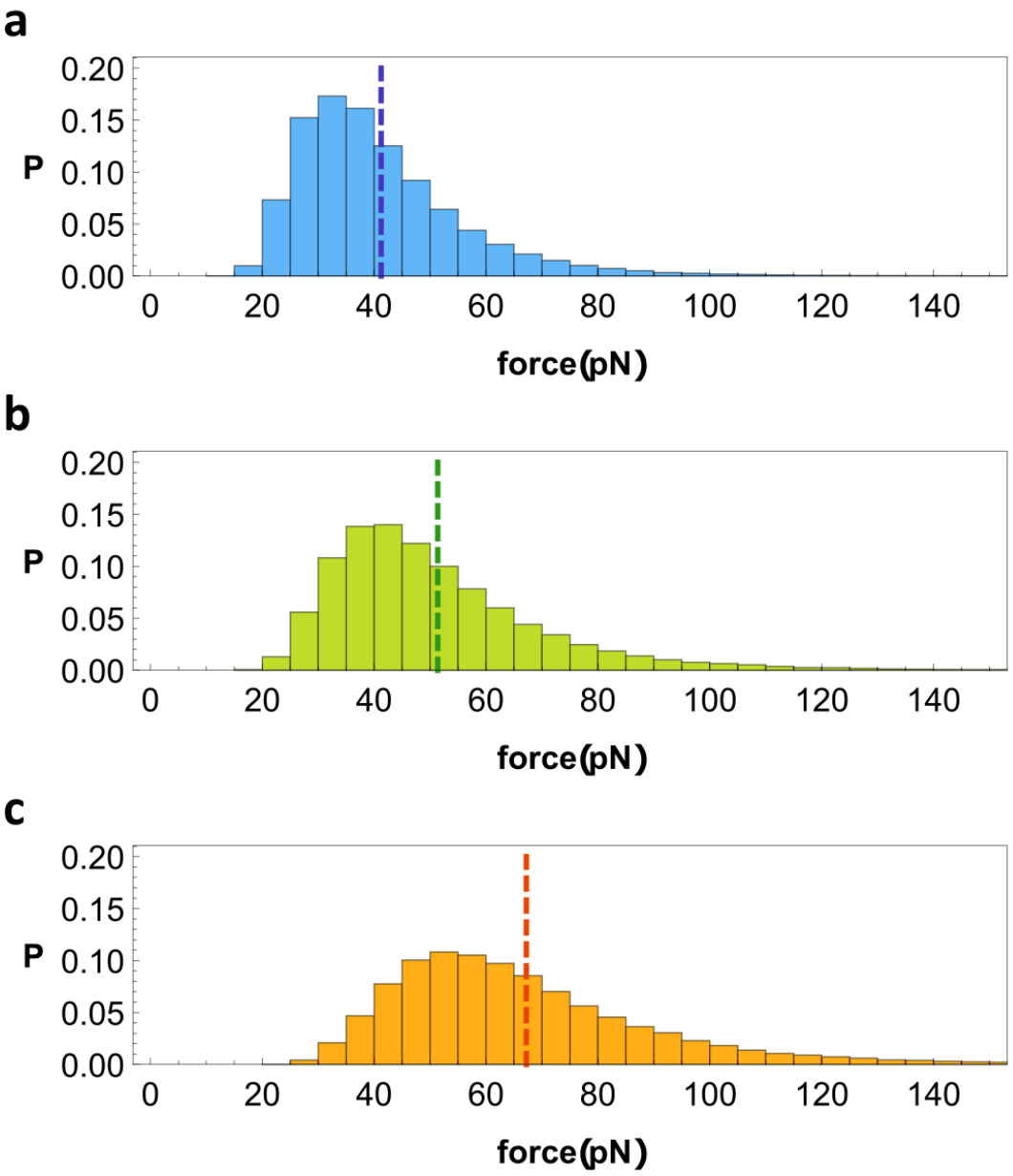


Figure 6.15: Probability distribution of the tension along the LH's NL, obtained through a time average during the waiting time between steps. The tension was estimated using the worm-like chain model (see appendix C). The histogram in a) blue corresponds to the unconstrained simulations. The histograms in b) green and c) orange correspond to the simulations of the first and second steps respectively where the rotation of the CC was constrained. The doted lines correspond to the mean values of the tension. These are $41.24 \pm 16.34pN$, $51.36 \pm 20.82pN$, and $67.20 \pm 25.83pN$ for the unconstrained simulations, first step simulations and second step simulations respectively.

## 6.3 Conclusion

Our simulations can track the behavior of kinesin at a resolution which is not typically accessible to experiments, giving us insight into the workings of the kinesin step. We are thus able to tie together the information available from a range of seemingly contradictory experimental results in a consistent manner, and paint a clear picture of the kinesin stepping mechanism at the molecular level.

In the matter of the symmetric versus asymmetric models of the kinesin step, it is our conclusion that, as is typical of molecular systems, the behavior of kinesin is stochastic and does not necessarily adhere to one particular mode of movement. In particular, we find that the presence of a cargo, bound to the kinesin motor through its CC, has a dramatic effect on the motor's stepping pattern.

When the motor is free of cargo, the TH has a preference for stepping through the right side of the LH. However, it is only slightly more probable for it to step twice consecutively on the same side than it is to alternate between left and right. Therefore, it is able to step in both a symmetric as well as an asymmetric fashion.

If the motor is bound to a large enough cargo we find evidence that its stepping pattern does tend to follow the asymmetric model. Nevertheless, even under these circumstances, we are unable to rule out the possibility that the motor occasionally steps in a symmetric manner.

Our most important conclusion is that the set up of the system has a significant effect on its behavior and that the method of measurement in an experiment may alter the results that are obtained. Whether the cargo indeed alters the stepping behavior of kinesin can be tested by either changing the size of the cargo or by tracking the individual MDs in the presence and absence of a cargo attached to the CC.

# Chapter 7

# Conclusion

## 7.1 Summary

One of the major challenges in biophysics is the correct interpretation of experimental results. When it comes to the study of molecular machines this challenge proves difficult as the spatial and temporal resolution of available technology can take us only so far. To mitigate this problem we use theoretical and computational modeling to gain insight into the workings of motor proteins. More specifically, CG models can be particularly useful when attempting to model the behavior of such systems[68,69,93,94]. Here we used two such approaches. The first was a mechanical toy model hybridized with a discrete state Markov chain, used to describe the stepping mechanism of dynein. The second was Brownian dynamic simulations of the SOP model[69,95,96,97]. This approach was applied to both dynein and kinesin.

Our analytical model for the stepping mechanism of dynein reproduces a wide variety of properties of the motor on both quantitative and qualitative levels. The most notable of these is the dynein step size distribution which, unlike other molecular motors, has an irregular shape and includes backward steps[13,14,39]. Our model reveals that the source of these backward steps is non-alternating stepping in which

the same MD steps twice in a row. Furthermore, we were able to derive a relatively simple analytical expression of the dwell time distribution at low ATP concentrations.

The advantage of analytical models such as the one we developed for the stepping mechanism of dynein is that they make it easy to explore a wide range of parameter space. When the model's parameters have concrete physical interpretations, as is the case in our model, we can make predications as to how dynein responds to changes in its physical environment. This becomes particularly useful when there is uncertainty as to the value of one of the parameters, as is the case for the ATP hydrolysis rate in dynein[18,40,43,46]. Using simple reasoning we were able to derive an expression for the physical constraints of the hydrolysis rate, power stroke rate, and the velocity of the motor.

Coarse grained simulations are also extremely successful in modeling the behavior of motor proteins. Here we use the SOP model to simulate motor proteins on two different functional levels. On the first level we explore the internal conformational changes in dynein and the mechanisms of allosteric communication in the motor. On the second level we study the stepping mechanics of kinesin by simulating the full motor dimer.

In the case of dynein we used a variant of the SOP model with double well energy potentials[68]. This allowed us to probe the spontaneous response of the dynein model to different external stimuli. While the number of dynein structures available has grown dramatically in recent years, these serve as snap shots of the system and do not provide direct information about its dynamics[17,18,19,43]. Nevertheless, when used as references for the SOP Hamiltonian, these structures enable us to explore the mechanisms behind allosteric communication in the MD. More specifically, the insert loops in the AAA2 domain of the dynein motor are known to play an important part in dynein activity, however, their precise function is not clear. Here we have shown

that the interactions between the AAA2 insert loops and the LN domain are crucial for the ATP induced priming stroke as well as the AAA3 controlled suppression mechanism. Furthermore, they serve as an external strain sensor in the motor's gating mechanism.

The SOP model is useful, not only in the context of internal conformational dynamics, but also in terms of the global stepping dynamics, as is the case of kinesin. Brownian dynamics simulations of the kinesin step, using hydrodynamic interactions (HI), have proven to be crucial for understanding the specifics of the kinesin stepping mechanism[69]. Here we use the kinesin SOP model to address the question of whether kinesin steps in a symmetrical or asymmetrical manner[84,87,88]. Our results show that the answer to this question depends on whether kinesin is bound to a large cargo or not. Furthermore, our simulations suggest that the kinesin step is a highly stochastic process and that the motor stepping mechanism is to some extent in the middle between the two extreme models. More importantly, perhaps, we show that experimental results that appear on the surface to be in contradiction are not necessarily so.

## 7.2 Future Perspectives

The success of the models presented here in addressing several important questions in the field of motor proteins suggests that similar approaches can be applied successfully to other motors. Our analytical model for dynein was based on a theory that has already been successfully applied to myosin V[67]. Myosin VI has an irregularly shaped step size distribution and similarly to dynein, it performs backward steps under zero backward load. This makes myosin VI a perfect target for a similar approach. In general, analytical models such as the one we used here are not only useful but also crucial in the analysis of motor proteins, especially in non trivial

cases such as dynein and myosin VI.

Similarly, there are open questions pertaining to dynein and kinesin that may be solved using CG models such as the SOP model. While the mechanisms with which information is passed along the dynein stalk to the MTBD is fairly well understood, the details of how affinity for MT is controlled are just emerging[50,51,52,98,99]. More specifically, how the application of external force at the LN domain translates into a conformational change in the MTBD and a change in MT affinity is still unclear. Coarse grained simulations of the dynein stalk and MTBD may prove to be extremely useful in understanding this process.

Regarding kinesin, the details of the stepping mechanism are reasonably well understood, however, the question of how allostery and gating work is still open[63,69,82,84]. Ironically, the fact that kinesin is significantly smaller than dynein makes it harder to investigate its allosteric mechanisms as they involve more subtle conformational changes. Nevertheless, the structural details of the different kinesin conformations are beginning to emerge[16]. Coarse grained models are particularly suited for such a scenario and could prove useful in understanding how kinesin responds to external stimuli, and more specifically, how internal strain between the two MDs translates into changes in MT affinity. We hope to pursue these different research directions in the near future.

# Appendix A

# Details of Analytical Model for Dynein

## A.1 CG Model from Crystal Structures

In order to arrive at any expressions for quantitatively describing the stepping of dynein we have to necessarily derive a CG model. A key concept in constructing such a model is to ensure that the CG description is consistent with the known structures of dynein. A basic unit in our model (Fig. 3.1) that plays an important role in the dynein mobility is the rod length.

The rod length, L, in each conformation as well as the orientation vector $\hat{\mathbf{u}}$ (Fig. 3.1) can be obtained using crystal structures of dynein in both the post- and pre-stroke conformations obtained from the Protein Data Bank (PDB) (PDBIDs: 3VKH,4RH7) as well as a structure of the MTBD bound to MT (PDBID: 3J1U) [18,19,21]. The vector connecting the LN domain and the MTBD is used to define these quantities. In the pre-stroke conformation (4RH7) we defined this vector by the positions of the $C_\alpha$ atoms in residues E1257 and E2980. In the post-stroke state (3VKH) the defining residues were E1524 and K3368. The length of this vector defines the size

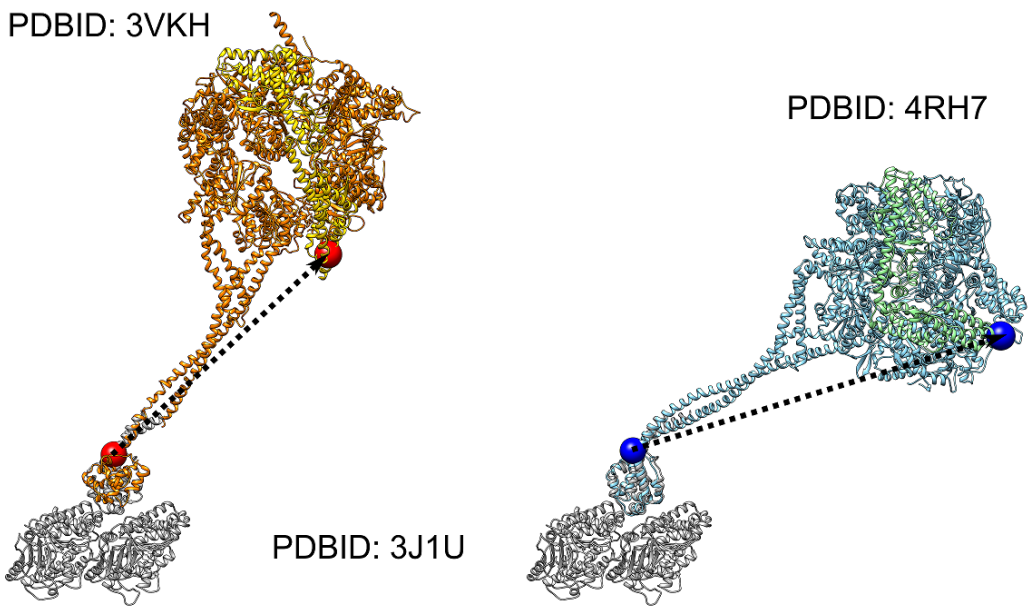


Figure A.1: Alignment of dynein crystal structures in the post stroke (3VKH) and pre stroke (4RH7) states with the MT bound MTBD structure (3J1U). The dashed arrows indicate the directions of $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$.

of L in each conformation. In order to obtain the orientation vector $\hat{\mathbf{u}}$, we calculated the orientation of the vector connecting the LN and the MTBD with respect to the MT axis. Since the structure containing the MTBD bound to the MT (3J1U) lacks the rest of the MD we had to align it with the full structure of the MD in each conformation (Fig. A.1). The pre-stroke structure (4RH7) was aligned by minimizing the Root Mean Squared Deviation (RMSD) between the corresponding $C_\alpha$ atoms in the MTBDs of both structures (4RH7,3J1U). In the post-stroke structure (3VKH) the MTBD is misaligned due to a missing fragment of the coiled-coil so the alignment was done using only the overlapping parts of the stalk in both structures (3VKH,3J1U). Once the structures were properly aligned, the orientation vector $\hat{\mathbf{u}}$ was easily obtained (Fig. A.1).

## A.2 Justification for Model

Our CG model is clearly a gross simplification of the MD structure, and yet, our calculated step size distribution agrees well with experiment. We give arguments, by way of justifying our approach, for the plausible reasons for the success of the model in capturing certain aspects of dynein mobility. The MD can undergo many conformational transitions, most of which involve allosteric communication between different sub-domains. The conformational transitions relevant to the step size, however, are the priming stroke and power stroke. Both of these transitions involve predominately the bending and straightening of the LN[18,19,41,46,49]. The basic hypothesis of our mechanical model is that the stepping pattern along the MT is determined by the differences in the LN conformations between the two MDs in the dimer. It is reasonable to make this assumption since both the MDs are connected through their LNs, and, therefore, conformational changes in the LN affect the relative positions of the MDs. Since the main effect of the LN transition is to alter the reach of the MD, the vector connecting the MTBD and the LN is a natural choice for describing the stepping of the motor. The power stroke in our model presents itself as a change in L from the value in the pre-stroke to the one in the post-stroke state. The direction $\hat{\mathbf{u}}$ changes as well as there is a difference in the orientation of the MD with respect to the MT between the two crystal structures. The transition between these two states occurs in a discrete manner, and is coupled to the rate parameter, $k_p$, describing both the power stroke and ADP release (see Fig. 3.2). These changes are sufficient to account for the different stepping patterns of the motor as well as the step size distribution.

The rest of the conformational changes are implicitly included in the chemomechanical cycle, which affect parameters such as the MT unbinding rate or the gating parameter $\gamma$. We should stress, however, that not all aspects of the movement of

the motor on MT can be described with the theory created here. The successes of this work and our previous study on myosin V show that such approaches, which produce analytic expressions capable of predicting the outcomes of experiments, are invaluable in understanding the structural basis of motor function.

An additional assumption we make for the mechanical model is that the system reaches equilibrium in terms of the MD orientation before the step is complete. In other words, the relaxation time for the rotational diffusion of a MD that is tethered to the MT is shorter than the typical time for ATP hydrolysis. To estimate of the rotational relaxation time, we treat the MD as a sphere with a diameter, $b = 10nm$ that is tethered to the MT at a constant distant, $R = 20nm$. The translational friction coefficient can be calculated using Stoke's formula,

$$\zeta = 3\pi\eta b \tag{A.1}$$

where $\eta = 0.01Poise$ is the viscosity of water. When the bead is tethered to the MT by a rigid constraint, the rotational friction coefficient is,

$$\zeta_r = R^2\zeta \tag{A.2}$$

The rotational relaxation time is given by,

$$\tau_r = \frac{1}{2D_r} = \frac{\zeta_r}{2K_BT} \approx 4.6\mu s \tag{A.3}$$

Since the hydrolysis time is on the order of milliseconds the system should reach equilibrium before stepping is complete, thus justifying the assumption.

## A.3 Derivation of the Mechanical Model

The equilibrium probability distribution for a stiff rod pointing in a random direction is given by,

$$\mathcal{P}(\mathbf{r}) = \frac{\delta(r-L)}{4\pi L^2}, \tag{A.4}$$

where $\mathbf{r}$ represents the end-to-end vector of the rod and L is the associated length.

When bound to the MT, the MD assumes a particular orientation with respect to the MT axis. In order to account for this preferred orientation we introduce an angular constraint in the form of a unit vector pointing towards the preferred direction. The normalized probability distribution $\int d\mathbf{r}\mathcal{P}(\mathbf{r}) = 1$, in the presence of the angular constraint is,

$$\mathcal{P}(\mathbf{r}) = \frac{\mathcal{T}}{\sinh\mathcal{T}}\frac{\delta(r-L)}{4\pi L^2}\exp(\mathcal{T}\hat{\mathbf{u}}\cdot\hat{\mathbf{r}}), \tag{A.5}$$

where $\hat{\mathbf{u}}$ is the unit vector pointing along the angular constraint and $\mathcal{T}$ is the parameter that determines the strength of the constraint.

The two MDs in the dimer are connected at their LN N-termini by means of complementary domains, capable of forming a dimer. We assume that the contact point between the two LNs is flexible. Therefore we treat the joint that connects the two rods in our model as a freely rotating joint.

In the case of a dimer of MDs, the vector connecting the two MTBDs is $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2$. The equilibrium probability distribution for the end-to-end vector can be calculated using,

$$\mathcal{P}(\mathbf{R}) = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \mathcal{P}(\mathbf{r}_1)\mathcal{P}(\mathbf{r}_2)\delta(\mathbf{R}-\mathbf{r}_1+\mathbf{r}_2). \tag{A.6}$$

By substituting Eq. A.5 for $\mathcal{P}(\mathbf{r})$ we obtain,

$$\mathcal{P}(\mathbf{R}) = N \int d\mathbf{r}_1 \int d\mathbf{r}_2 \delta(r_1 - L_1)\,\delta(r_2 - L_2) \exp\left(\mathcal{T}_1 \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{r}}_1 + \mathcal{T}_2 \hat{\mathbf{u}}_2 \cdot \hat{\mathbf{r}}_2\right) \delta(\mathbf{R} - \mathbf{r}_1 + \mathbf{r}_2) \tag{A.7}$$

where $N = \frac{\mathcal{T}_1}{\sinh \mathcal{T}_1} \frac{\mathcal{T}_2}{\sinh \mathcal{T}_2} \frac{1}{16\pi^2 L_1^2 L_2^2}$ is the normalization constant. After integrating over $\mathbf{r}_2$ we obtain,

$$\mathcal{P}(\mathbf{R}) = N \int d\mathbf{r}_1 \delta(r_1 - L_1)\,\delta(|\mathbf{r}_1 - \mathbf{R}| - L_2) \exp\left(\mathcal{T}_1 \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{r}}_1 + \mathcal{T}_2 \hat{\mathbf{u}}_2 \cdot \frac{\mathbf{r}_1 - \mathbf{R}}{|\mathbf{r}_1 - \mathbf{R}|}\right) \tag{A.8}$$

The integral in Eq. A.8 is evaluated in spherical coordinates. Integrating over the radial components fixes the value of $r_1$ at $L_1$. Since $|\mathbf{r}_1 - \mathbf{R}|$ is just $r_2$, it can be replaced by $L_2$. Furthermore, by fixing the three lengths $R$, $r_1$, and $r_2$, the value of $\theta_1$ is constrained, thus reducing Eq. A.8 to a one dimensional integral,

$$\mathcal{P}(\mathbf{R}) = N \int L_1^2 \sin\theta_1 d\phi_1 \exp\left(\mathcal{T}_1 \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{r}}_1 + \mathcal{T}_2 \hat{\mathbf{u}}_2 \cdot \frac{\mathbf{r}_1 - \mathbf{R}}{L_2}\right). \tag{A.9}$$

By taking all the elements that do not depend on $\phi_1$ out of the integral and recognizing that $\mathbf{r}_1 = L_1 \hat{\mathbf{r}}_1$ we obtain,

$$\mathcal{P}(\mathbf{R}) = N L_1^2 \sin\theta_1 \exp\left(-\mathcal{T}_2 \hat{\mathbf{u}}_2 \cdot \frac{\mathbf{R}}{L_2}\right) \int d\phi_1 \exp\left(\left(\mathcal{T}_1 \hat{\mathbf{u}}_1 + \mathcal{T}_2 \frac{L_1}{L_2} \hat{\mathbf{u}}_2\right) \cdot \hat{\mathbf{r}}_1\right) \tag{A.10}$$

We define a new constraint vector as a linear combination of the two constraints,

$$\begin{aligned} \mathbf{u}_c &= \mathcal{T}_1 \hat{\mathbf{u}}_1 + \mathcal{T}_2 \frac{L_1}{L_2} \hat{\mathbf{u}}_2 \\ \mathcal{T}_c &= |\mathbf{u}_c| \end{aligned} \tag{A.11}$$

and rewrite our integral as,

$$\mathcal{P}(\mathbf{R}) = N L_1^2 \sin\theta_1 \exp\left(-\mathcal{T}_2 \hat{\mathbf{u}}_2 \cdot \frac{\mathbf{R}}{L_2}\right) \int d\phi_1 \exp\left(\mathbf{u}_c \cdot \hat{\mathbf{r}}_1\right) \tag{A.12}$$

The following dot product can be written as:

$$\hat{\mathbf{u}}_c \cdot \hat{\mathbf{r}}_1 = \cos\theta_1 \cos\theta_c + \cos\phi_1 \sin\theta_1 \sin\theta_c \tag{A.13}$$

where $\theta_c$ is the angle between $\mathbf{u}_c$ and $\mathbf{R}$. We can then perform the integration, resulting in the final expression for $\mathcal{P}(\mathbf{R})$,

$$\begin{gathered}\mathcal{P}(\mathbf{R}) = 2\pi N L_1^2 \sin\theta_1 \exp\left(-\mathcal{T}_2 \hat{\mathbf{u}}_2 \cdot \frac{\mathbf{R}}{L_2} + \mathcal{T}_c \cos\theta_c \cos\theta_1\right) \\ \times \\ I_0\left(\mathcal{T}_c \sin\theta_c \sin\theta_1\right)\end{gathered} \tag{A.14}$$

where $I_0$ is the modified Bessel function of the first kind with order zero, $\theta_1$ is the angle between $\mathbf{R}$ and $\mathbf{r}_1$, and $\theta_c$ is the angle between a linear combination of the constraint vectors and $\mathbf{R}$.

By using the law of cosines and dot product we obtain the following expressions for the trigonometric elements in our function:

$$\begin{gathered}\cos\theta_1 = \frac{R^2 + L_1^2 - L_2^2}{2RL_1} \qquad\qquad \cos\theta_c = \hat{\mathbf{u}}_c \cdot \hat{\mathbf{R}} \\ \sin\theta_1 = \sqrt{1 - \left(\frac{R^2 + L_1^2 - L_2^2}{2RL_1}\right)^2} \qquad \sin\theta_c = \sqrt{1 - \left(\hat{\mathbf{u}}_c \cdot \hat{\mathbf{R}}\right)^2}\end{gathered} \tag{A.15}$$

It is important to note that the constraint vector $\hat{\mathbf{u}}_2$ is a function of the binding site and is therefore a function of $\mathbf{R}$. To calculate its direction we simply rotate our original vector according to the angle between the line connecting the binding site to the center of the MT and the $\hat{\mathbf{z}}$ axis.

## A.4 Discrete to Continuum Step Size Distribution

The step size distributions obtained using Eq. 3.5 and Eq. 3.7 are discrete. This is because that there is only a finite number of TBS, each at a fixed position. In contrast, the step size distributions obtained from single molecule experiments are continuous. The reason for this is an intrinsic error in the measurement of the position of the quantum dots labeling the motor. In order to account for this "noise" and for the lack of perfect registry between the protofilaments in the MT, we convolve our discrete distribution with a Gaussian function,

$$\mathcal{P}\left(\Delta x\right)_{cont} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{P}\left(\Delta y\right) \exp\left(-\frac{\left(\Delta x - \Delta y\right)^2}{2\sigma^2}\right) d\Delta y \tag{A.16}$$

where $\sigma$ has been chosen to be $3nm$ which is comparable with the typical scale of measurement error in the experiments. Furthermore, since experiments cannot detect steps with $\Delta x < 3nm$, we did not include steps below 3nm. However, the inclusion of smaller steps was necessary for the purpose of calculating the mean step size and the velocity.

## A.5 Markov Chain Model Derivation

In order to determine the transition probabilities between the different stepping states we used the kinetic scheme described in Fig. 3.4 and solved the first passage time problem for each transition using the method employed by Shaevitz et.al[100]. In addition, we obtained the corresponding dwell time distributions and their moments. Due to the necessarily complex nature of the kinetic model, these expressions are

lengthy. Therefore, we solved them numerically using Mathematica 10 without writing the analytical formulas explicitly[101].

Using these transition probabilities, we initially constructed a seven state model, $S'_n = \{A_1, B_1, C_1, A_2, B_2, C_2, T_{er}\}$, which includes the termination state. This model allows for the calculation of the run termination probability. The matrix,

$$M'_{i,j} = \mathcal{P}\left(S_{n+1}\left(j\right)|S_n\left(i\right)\right) \tag{A.17}$$

is the corresponding $7 \times 7$ Markov matrix whose elements are the transition probabilities between the stepping states.

In order to find the stationary probability of stepping in each mode we make the following approximation. As long as the termination probability is sufficiently small we can assume that the system reaches a pseudo steady state. In such a scenario, we can construct a $6 \times 6$ Markov matrix, $M$, whose elements are the renormalized transition probabilities between the stepping states, excluding the termination state $T_{er}$. The stationary probability distribution, $\mathcal{P}\left(S_i\right)$, is then obtained by calculating the eigenvector of $M$ that corresponds to the first eigenvalue.

We now turn to the calculation of the dwell times. Using the mechanical model we can calculate the position of the individual MD but not of the tail domain (the joint connecting the two MDs). Therefore, we are interested in calculating the dwell times between steps of the same MD and not of the motor dimer as a whole. We can also calculate the probability of a single MD stepping twice consecutively, $P_c$, as well as the probability of an alternating step, $P_a$, using,

$$P_c = \sum_{i=1}^{3}\sum_{j=1}^{3}\mathcal{P}\left(S_i\right)M_{i,j} + \sum_{i=4}^{6}\sum_{j=4}^{6}\mathcal{P}\left(S_i\right)M_{i,j}, \tag{A.18}$$

$$P_a = \sum_{i=1}^{3}\sum_{j=4}^{6}\mathcal{P}\left(S_i\right)M_{i,j} + \sum_{i=4}^{6}\sum_{j=1}^{3}\mathcal{P}\left(S_i\right)M_{i,j}. \tag{A.19}$$

Similarly, the mean dwell time, $\tau_c$, between two consecutive steps of the same MD and the mean dwell time between two alternating steps, $\tau_a$, are given by:

$$\tau_c = \sum_{i=1}^{3}\sum_{j=1}^{3} \mathcal{P}(S_i) M_{i,j} T_{i,j} + \sum_{i=4}^{6}\sum_{j=4}^{6} \mathcal{P}(S_i) M_{i,j} T_{i,j}, \tag{A.20}$$

$$\tau_a = \sum_{i=1}^{3}\sum_{j=4}^{6} \mathcal{P}(S_i) M_{i,j} T_{i,j} + \sum_{i=4}^{6}\sum_{j=1}^{3} \mathcal{P}(S_i) M_{i,j} T_{i,j}. \tag{A.21}$$

The mean dwell time between steps of a single MD can be calculated by considering all possible scenarios such as two consecutive steps of the same MD, two alternating steps, and so forth. The dwell time is then given by the infinite sum:

$$\begin{aligned} \tau_{MD} &= P_c\tau_c + P_a^2(2\tau_a) + P_a^2 P_c(2\tau_a + \tau_c)... \\ &= P_c\tau_c + P_a^2 \textstyle\sum_{n=0}^{\infty} P_c^n (2\tau_a + n\tau_c). \end{aligned} \tag{A.22}$$

which converges to,

$$\tau_{MD} = P_c\tau_c + \frac{2\tau_a + P_c\tau_c - 2\tau_a P_c}{(P_c - 1)^2}. \tag{A.23}$$

In order to solve for the dwell time distribution one has to replace the infinite sum of the mean dwell times with an infinite number of convolutions of the dwell time distribution functions, $f_c(t)$ and $f_a(t)$, which makes the calculation extremely complicated. Nevertheless, we can overcome this obstacle by using the Laplace transform, $\tilde{f}(s) = \int_0^\infty dt e^{-st} f(t)$, where convolution integrals can be treated as products. We then obtain:

$$f_{MD}(t) = P_c\tau_c + \mathcal{L}^{-1}\left(P_a^2 \tilde{f}_a(s)^2 \sum_{n=0}^{\infty} P_c^n \tilde{f}_c(s)^n\right) \tag{A.24}$$

This product also converges, leading to the result:

$$f_{MD}(t) = P_c\tau_c + \mathcal{L}^{-1}\left(\frac{P_a^2\tilde{f}_a(s)^2}{1 - P_c\tilde{f}_c(s)}\right) \tag{A.25}$$

## A.6 Parameter Fitting Procedure

There are four parameters in our model, whose values we do not know a priori. They are the ATP binding rate $k_{ATP}$, the tension parameter $\mathcal{T}$, the gating parameter $\gamma$, and the power stroke rate $k_p$. In order to obtain their values we fit simultaneously four observables to experiment. The observables are the velocity, run length at saturating ATP concentrations, the mean step size, and the mean dwell time at low ATP concentrations[13,14].

The fitting was done by calculating the goodness of fit,

$$\chi^2 = \frac{(vel - vel_{exp})^2}{S.E.(vel)^2} + \frac{(run - run_{exp})^2}{S.E.(run)^2} + \frac{(mdt - mdt_{exp})^2}{S.E.(mdt)^2} + \frac{(mss - mss_{exp})^2}{S.E.(mss)^2} \tag{A.26}$$

where $vel$, $run$, $mdt$, and $mss$ stand for velocity, run length, mean dwell time and mean step size respectively. The subscript $exp$ denotes the experimentally determined value and S.E. is the standard error. We minimized $\chi^2$ with respected to the four parameters using the gradient descent approach. The results are given in Table S1.

## A.7 Approximate Expression for the Dwell Time Distribution

Since the expression for the dwell time distribution, given by Eq. A.25, is complicated, we are interested in finding a simpler expression that approximates the result

| Obs. | Calc. Value | Exp. Mean | Exp. S.D. | Exp. N | Exp. S.E. | Source |
|---|---|---|---|---|---|---|
| vel | $\approx 134nm/s$ | $134nm/s$ | $60.4nm/s$ | 943 | $\approx 3.9nm/s$ | [14] |
| run | $\approx 1.06\mu m$ | $1.06\mu m$ | - | - | $0.044\mu m$ | [14] |
| mdt | $\approx 0.599s$ | $0.599s$ | $0.546s$ | 552 | $\approx 5.4 \times 10^{-4}s$ | [13] |
| mss | $\approx 9.55nm$ | $10.2nm$ | $17.1nm$ | 734 | $\approx 0.4nm$ | [13] |
| $\chi^2$ | 0.53 | | | | | |

Table A.1: Comparison between calculated and experimental values of observables. The list of observables is velocity (vel), run length (run), mean dwell time (mdt), and mean step size (mss).

and can be easily fit to experimental data. Since the dwell time measurements are typically done under low ATP concentration, we assume that the binding of ATP is rate limiting in the process. In addition, if we assume that both MT detachment and ATP hydrolysis are fast compared to ATP binding the power-stroke rate, an assumption that is correct with our choice of parameters, we can approximate our kinetic scheme using the scheme depicted in Fig. 3.7. The solution to Eq. A.25 is given by:

$$k_{ATP}k_p(-\frac{e^{t\left(-0.5\sqrt{4k_{ATP}^2+k_p^2}-k_{ATP}-0.5k_p\right)}}{\sqrt{4k_{ATP}^2+k_p^2}}+\frac{e^{t\left(0.5\sqrt{4k_{ATP}^2+k_p^2}-k_{ATP}-0.5k_p\right)}}{\sqrt{4k_{ATP}^2+k_p^2}} \\ -\frac{e^{-2k_{ATP}t}}{2k_{ATP}-k_p}+\frac{0.5e^{-2k_{ATP}t}}{k_{ATP}-0.5k_p}-\frac{e^{-k_pt}}{k_p-2k_{ATP}}+\frac{e^{-k_pt}}{k_p-2k_{ATP}}) \tag{A.27}$$

However, because $k_p^2 >> k_{ATP}^2$, we can rewrite $\sqrt{4k_{ATP}^2+k_p^2}$ as $k_p$. Once we make the substitution and re-normalize Eq. A.27, we obtain,

$$\tau_{MD} = \frac{e^{-k_{ATP}t}\left(1-e^{-k_pt}\right)k_{ATP}\left(k_{ATP}+k_p\right)}{k_p}. \tag{A.28}$$

# Appendix B

# Details of SOP Model - Dynein

## B.1 Reference Structures

Our CG model relies on the existence of known structures in order to construct distance maps that serve as reference for the Hamiltonian. Thankfully, several crystal structures of dynein at different conformations and nucleotide states have been discovered in recent years[17,18,19,43]. In this study we used as reference four different crystal structures which are available in the PDB: (1) Human dynein in the pre power stroke conformation, with an ATP analog bound to the AAA1/2 binding site (4RH7). (2) Slime mold dynein in the post power stroke conformation with ADP bound to the AAA1/2 site (3VKG). (3) Yeast Dynein in the post power stroke state with no nucleotide bound to the AAA1/2 site (4AKG). (4) Yeast dynein in the post stroke conformation with an ATP analog bound to the AAA1/2 site as well as the AAA3 site (4W8F).

In order to construct compatible distance maps from crystal structures of different species, we needed to first build a proper sequence alignment. We first disassembled each structure into its sub-domains. We then aligned all the pairs of corresponding sub-domains using the VMD implementation of the MultiSeq pack-

| Domain | Subunit | 4RH7 | 3VKG | 4AKG/4W8F | Model |
|---|---|---|---|---|---|
| Linker | A | Q1255-F1426 | Q1522-F1694 | F1365-F1534 | 0-171 |
| | B | P1427-T1665 | A1695-T1950 | P1535-T1772 | 172-410 |
| AAA1 | Large | P1666-S1818 | P1951-I2101 | P1773-K1922 | 411-563 |
| | Small | H1819-I1955 | K2102-K2252 | S1923-S2050 | 564-700 |
| AAA2 | Large | P1956-E2119 | Q2253-I2420 | E2051-V2219 | 701-864 |
| | Small | T2120-T2265 | L2421-T2650 | C2220-T2394 | 865-1010 |
| AAA3 | Large | P2266-Y2432 | V2651-F2818 | I2395-Y2561 | 1011-1177 |
| | Small | P2433-F2630 | P2819-F2991 | P2562-H2735 | 1178-1372 |
| AAA4 | Large | H2631-E2809 | N2992-G3174 | E2736-G2918 | 1373-1551 |
| | Small+Stalk | G2810-D3261 | E3175-D3655 | D2919-N3387 | 1552-1983 |
| AAA5 | Large | D3262-T3393 | E3656-T3787 | D3388-V3519 | 1984-2115 |
| | Strut | T3394-L3563 | V3788-L3959 | T3520-L3690 | 2116-2285 |
| | Small | F3564-N3709 | L3960-D4117 | D3691-N3373 | 2286-2415 |
| AAA6 | Large | L3710-E3831 | M4118-E4239 | I3774-E3898 | 2416-2537 |
| | Small | S3832-P3999 | N4240-P4411 | D3899-P4066 | 2538-2695 |

Table B.1: Dynein domains sequence positions with numbers showing the start and end of each domain.

age[102]. Since 4RH7 is the most complete structure for dynein, we used it as the reference structure in our alignments. Table B.1 shows the sequence boundaries of each domain and subunits as well as the corresponding sequence position in our model.

Using our alignment and the coordinates obtained from the crystal structures, we built four distance maps, each corresponding to a different crystal structure. Inter-residue distances were measured as the distance between the corresponding $C_\alpha$ carbons. In the cases in which there were gaps in the structure or misalignments, the 4RH7 distances were used as a default. The 4RH7 structure itself does contain a small number of gaps and the residues in those gaps were left unresolved in our model.

## B.2 SOP Model

CG models have been remarkably successful when used in simulations of biological molecules[68,69,93,94]. Such models are particularly useful when the system in question is composed of large proteins as is the case in dynein. Here, we use a variation on the SOP model with double well potential[68]. Each amino acid in our model is represented by a single bead, positioned at the location of the $C_\alpha$ carbon. The positions of the beads are obtained from the reference crystal structures. The Hamiltonian of the system is:

$$
\begin{aligned}
U\left(\{\bar{r}_i\}\right) &= U_{FENE} + U_{nb}^{(rep)} + U_{nb}^{(\alpha)} + U_{nb}^{(\alpha\beta)} + U_{nb}^{(\beta)} + U_{nb}^{(x)} \\
U_{FENE} &= -\sum_{i=1}^{N-1} \frac{k}{2} R_0^2 \log\left(1 - \frac{\left(r_{i,i+1} - r_{i,i+1}^0\right)^2}{R_0^2}\right) \\
U_{nb}^{(\alpha\beta)} &= \sum_{i=1}^{N-3}\sum_{i=i+3}^{N} \epsilon_h \min\left\{\left[\left(\frac{r_{ij}^{0\alpha}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{0\alpha}}{r_{ij}}\right)^{6}\right], \lambda_{\alpha\beta}\left[\left(\frac{r_{ij}^{0\beta}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{0\beta}}{r_{ij}}\right)^{6}\right]\right\} \Delta_{ij}^{\alpha}\Delta_{ij}^{\beta} \\
U_{nb}^{(\alpha)} &= \sum_{i=1}^{N-3}\sum_{i=i+3}^{N} \epsilon_h \left[\left(\frac{r_{ij}^{0\alpha}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{0\alpha}}{r_{ij}}\right)^{6}\right] \Delta_{ij}^{\alpha}\left(1 - \Delta_{ij}^{\beta}\right) \\
U_{nb}^{(\beta)} &= \sum_{i=1}^{N-3}\sum_{i=i+3}^{N} \epsilon_h \left[\left[\left(\frac{r_{ij}^{0\beta}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{0\beta}}{r_{ij}}\right)^{6}\right]\left(1 - H\left(r_{ij} - r_{ij}^0\right)\right)\right. \\
&\quad \left. + \left[\lambda_{\alpha\beta}\left(\left(\frac{r_{ij}^{0\beta}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{0\beta}}{r_{ij}}\right)^{6}\right) + \lambda_{\alpha\beta} - 1\right] H\left(r_{ij} - r_{ij}^0\right)\right]\left(1 - \Delta_{ij}^{\alpha}\right)\Delta_{ij}^{\beta} \\
U_{nb}^{(rep)} &= \left[\sum_{i=1}^{N-2} \epsilon_l \left(\frac{\sigma}{r_{i,i+2}}\right)^6 + \sum_{i=1}^{N-3}\sum_{i=i+3}^{N} \epsilon_l \left(\frac{\sigma}{r_{ij}}\right)^6 \left(1 - \Delta_{ij}^{\alpha}\right)\left(1 - \Delta_{ij}^{\beta}\right)\right]
\end{aligned}
\tag{B.1}
$$

The first term in the Hamiltonian accounts for chain connectivity, given by the Finitely Elastic Nonlinear Elastic (FENE) potential, where $r_{i,i+1}$ is the distance between beads i and i+1 and $r_{i,i+1}^0$ is the corresponding distance in the reference

distance map. Terms 2 through 4 represent stabilizing interactions for non-bonded residues which are found to be in contact in the reference crystal structure. A pair of residues is defined as a contact if the distance between them is equal to or smaller than a cutoff, $r_c = 10\mathring{A}$. The last term accounts for non bonded volume exclusion interactions between residues that are not in contact in either reference structure.

Dynein is thought to be found predominantly in one of two conformational states: the post stroke apo state (referred to as state $\alpha$) and the ATP bound pre stroke state (referred to as state $\beta$). Since dynein can alternate between these two conformations, we use a double well potential for those residues which are found to be in contact in both reference states ($U_{nb}^{(\alpha\beta)}$). $\lambda_{\alpha\beta}$ is a dimensionless parameter with a value between 0 and 1 that determines the stability of the $\beta$ state relative to $\alpha$. When dynein is not bound to any nucleotides, it is found in the $\alpha$ state. We therefore use $\alpha$ as the default state of the system and assume that it is more stable in comparison to state $\beta$. To reflect this observation, we set the value of $\lambda_{\alpha\beta}$ to 0.5 in the Hamiltonian.

For those contacts that exist in state $\alpha$ only we use the Lennard-Jones potential ($U_{nb}^{(\alpha)}$). For contacts that are found only in the $\beta$ state we use a modification of the Lennard-Jones potential, $U_{nb}^{(\beta)}$, where $H$ is the unit step function (Eq. B.1). In this potential, the stabilizing region is modified by $\lambda_{\alpha\beta}$ to account for the fact that state $\beta$ is less stable in comparison to state $\alpha$.

## B.3 Construction of $\alpha$ and $\beta$ Distance Maps

The $\alpha$ state corresponds to the post power stroke conformation in the apo state. We therefore used the 4AKG crystal structures in order to generate the reference distances in the $\alpha$ distance map. Since most of the stalk domain is unresolved in the 4AKG crystal structure, we used the distances from the 3VKG structure for the

segments ranging from the AAA4 small subunit through the AAA5 small subunit. The distance map for the $\beta$ state was constructed using the 4RH7 structure, which corresponds to the pre power stroke, AAA1/2 ATP bound state.

In order to introduce the effects the AAA3 domain binding ATP and repressing dynein activity we used the distance from structure 4W8F in map $\alpha$ for segments ranging from the LN through the AAA4 large subunit. Interactions between the LN and the AAA2 subunits were also taken from the 4W8F structure and added to both $\alpha$ and $\beta$ maps to account for the fact that these interactions should be allowed regardless of the nucleotide state of the AAA3 unit.

## B.4 AAA1/2 ATP Binding Site Interactions

In the apo state we used the distances from 4AKG in both $\alpha$ and $\beta$ states for the AAA1/2 binding site (segments ranging from the LN B subunit through the AAA2 small subunit). In order to simulate ATP binding we used the distances from 4RH7 in the $\alpha$ map within the same segments. For those AAA1/2 interactions between the AAA1 and AAA2 domains, corresponding to contacts that exist only in the 4RH7 we used the following potential,

$$\begin{aligned} U_{nb}^{(ATP)} = \sum_{i=1}^{N-3} \sum_{i=i+3}^{N} \epsilon_h \Bigg[ \left[ \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^{6} \right] \left( 1 - H\left( r_{ij} - r_{ij}^0 \right) \right) \\ + \left[ \lambda_{AAA1/2} \left( \left( \frac{\sigma}{r_{ij} + \sigma - r_{ij}^0} \right)^{12} - 2 \left( \frac{\sigma}{r_{ij} + \sigma - r_{ij}^0} \right)^{6} \right) + \lambda_{AAA1/2} - 1 \right] \\ \times H\left( r_{ij} - r_{ij}^0 \right) \Bigg] \Delta_{ij}^{ATP}. \end{aligned} \tag{B.2}$$

This function is similar to the modified Lennard-Jones potential, $U_{nb}^{(\beta)}$. However, it has been further modified in order to keep a constant width of the energy

well. The purpose of this modification is to prevent the stabilizing interactions within the AAA1/2 cleft from having too large of a range. The dimensionless parameter $\lambda_{AAA1/2}$ determines the strength of the interactions. The interface for these interactions was defined as any contact between any residue in the LN B and AAA1 large subunits, and the AAA1 small subunit and AAA2 domain. In the simulations in which the system was in the ADP bound state, $\lambda_{AAA1/2}$ was set to half its value in the ATP bound state to account for the fact that ADP binding does not induce a transition to the pre power stroke state.

When ATP is bound to the AAA3 domain and the motor is in the repressed state, the interactions between the AAA1 and AAA2 domains within the AAA1/2 cleft, as seen in structure 4W8F, are different from the AAA1/2 interactions in the active ATP bound structure, 4RH7. Here we assume that these interactions are not native and are an indirect result of ATP binding to the AAA3 domain. We account for this by setting the strength of these interactions to 0.

## B.5 Linker (LN) Interactions

The LN domain remains bound to the AAA5 domain when the AAA1/2 site is in the apo state but detaches when ATP binds to the AAA1/2 site. Similarly, the LN domain remains bound to the AA3 domain in the ATP bound state but detaches once ADP has been released. This behavior depends on the strength of the interactions of the LN domain with the AAA3 and AAA5 domains as well as the interactions with the AAA2 ILs. To ensure this behavior is represented by our model we modified the interaction between the LN domain and the AAA2, AAA3, and AAA5 domains using the following potential function,

$$U_{nb}^{(x)} = \sum_{i=1}^{N-3} \sum_{i=i+3}^{N} \epsilon_h \left[ \left[ \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^{6} \right] \left( 1 - H\left( r_{ij} - r_{ij}^0 \right) \right) + \left[ \lambda_x \left( \left( \frac{\sigma}{r_{ij} + \sigma - r_{ij}^0} \right)^{12} - 2 \left( \frac{\sigma}{r_{ij} + \sigma - r_{ij}^0} \right)^{6} \right) + \lambda_x - 1 \right] H\left( r_{ij} - r_{ij}^0 \right) \right] \Delta_{ij}^x. \tag{B.3}$$

where x stands for one of the three LN interaction interfaces, AAA2, AAA3, and AAA5. Similarly, the parameters $\lambda_{LNK-AAA2}$, $\lambda_{LNK-AAA3}$, and $\lambda_{LNK-AAA5}$ determine the strength of the interaction with each binding site.

When the LN domain is in the pre power stroke conformation there are contacts that form between the A and B LN subunits. To ensure that these interactions do not prevent the LN domain from extending to perform the power stroke, we use the same potential function, $U_{nb}^{(x)}$, and modify their strength with $\lambda_{LNK}$. The $\lambda$ values as well as the values of the rest of the parameters are given in table B.2.

## B.6 Equations of Motion

We simulated the dynein motor using a numerical solution of Brownian dynamics[103]. The equation of motion is given by,

$$\bar{r}_i\left(t+h\right) = \bar{r}_i\left(t\right) + \frac{h}{\zeta}\left( -\frac{\partial U\left(\{\bar{r}_i\}\right)}{\partial \bar{r}_i} \right) + \bar{\Gamma}_i\left(t\right) \tag{B.4}$$

where $\bar{r}_i\left(t\right)$ is the position of bead i at time t, $h$ is the time interval between frames, $\zeta$ is the friction coefficient, and $\bar{\Gamma}_i\left(t\right)$ is a normally distributed random noise, acting on bead i, that accounts for thermal fluctuation. We require that $\bar{\Gamma}_i\left(t\right)$ obey,

$$\langle \mathbf{\Gamma}_i(t) \mathbf{\Gamma}_j(t') \rangle = 6 \frac{K_B T}{\zeta} h \delta_{ij} \delta_{tt'}. \tag{B.5}$$

| Parameter | Value |
|---|---|
| k | 20 $kcal/(mol\mathring{A}^2)$ |
| $R_0$ | 2Å |
| $\sigma$ | 3.8Å |
| $\epsilon_l$ | 1 $kcal/mol$ |
| $\epsilon_h$ | 2 $kcal/mol$ |
| $\lambda_{\alpha\beta}$ | 0.5 |
| $\lambda_{AAA1/2}$ | 0.1 |
| $\lambda_{LNK}$ | 0.07 |
| $\lambda_{LNK-AAA2}$ | 0.39 |
| $\lambda_{LNK-AAA3}$ | 0.25 |
| $\lambda_{LNK-AAA5}$ | 0.14 |

Table B.2: List of parameter values.

where $K_B$ is the Boltzmann constant and T is the temperature. The friction coefficient was determined using the Stokes equation, $\zeta = 6\pi\eta a$, where $\eta$ is the viscosity and $a = 1.9\mathring{A}$ is the bead radius.

## B.7 Time Scales

In order to obtain a numerically stable solution to the equation of motion, the time step has to be small enough. The upper bound is determined by the typical length scale of the stiffest term in the Hamiltonian, in this case the FENE potential. The typical length scale of the FENE potential can be approximated by $a \approx \sqrt{\frac{k_B T}{k}}$. The typical time scale for Brownian dynamics is given by $\tau = \frac{\zeta a^2}{k_B T}$. Since the bond lengths between sequence adjacent residues are approximately constant, the dynamics are dominated by the Lennard-Jones potential with a typical length scale $\sigma$. The time step can therefore be modified by the factor $\frac{\tau_{LJ}}{\tau_{FENE}} = \sigma^2 / \frac{k_B T}{k} \approx 500$

to yield $h = 45ps$.

## B.8 Switching Between States of the AAA5/6/s Domains

In simulations of the ATP induced transition between the post and pre power stroke, the Hamiltonian of the AAA5/6/s domains is switched from the M state to the U state, once $\Delta_{N1714-S2065}$ reaches a value smaller than 8Å. AAA5/6/s interactions are defined as interactions within the range from the AAA4 small subunit through the AAA6 domain, and interactions between the AAA6 domain, and the LN and AAA1 domains. Instead of using the double well potential for shared contacts, only one of the two reference maps, $\alpha$ or $\beta$, was used with a standard Lennard-Jones potential. The switch between states M and U is done by using the following transition function:

$$r_{ij}^{0}\left(\alpha \rightarrow \beta\right) = \left(1-k\right) r_{ij}^{0\alpha} + k r_{ij}^{0\beta}. \tag{B.6}$$

The reference distance $r_{ij}^{0}$ in the Lennard-Jones potential is transitioned from the distance in state $\alpha$ to the distance in state $\beta$ by gradually increasing the value of $k$ from 0 to 1 over a short period of time. Because the conformational change occurs over a long time the use of Eq. B.6 will not affect our results.

# Appendix C

# Details of SOP Model - Kinesin

## C.1 CG Model

The molecular construct of the two MD, bound to the MT, is a massive system from the perspective of molecular simulations. Due to the large size of the system we use CG models, which have been have to shown to successfully model biological systems, including motor proteins, both quantitatively and qualitatively[68,69,89]. Here we employ the use of the SOP model which has been used before successfully to elucidate the kinematics of the kinesin step[69]. Each bead in our model, centered at the location of the $C_\alpha$ carbon, represents a single amino acid. The interactions between the beads are determined by the contacts found in the crystal structures of the system (see supplementary information in Zhang and Thirumalai (ZT) for details[69]). The Hamiltonian of the system is given by,

$$
\begin{aligned}
H\left(r_i|X\right) &= V_{FENE} + V_{NB}^{A} + V_{NB}^{R} + V_{NB}^{E} \\
&= -\sum_{Y}\sum_{i=1}^{N^Y-1}\frac{k}{2}R_0^2\log\left(1-\frac{\left(r_{i,i+1}-r_{i,i+1}^0\left(X\right)\right)^2}{R_0^2}\right) \\
&+ \sum_{Y,Y'}\sum_{i,j}\epsilon_h^{Y-Y'}\left[\left(\frac{r_{ij}^0\left(X\right)}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^0\left(X\right)}{r_{ij}}\right)^{6}\right]\Delta_{ij} \\
&+ \sum_{Y}\sum_{i=1}^{N^Y-2}\epsilon_l\left(\frac{\sigma}{r_{i,i+2}}\right)^6 + \sum_{Y,Y'}\sum_{i,j}\epsilon_l\left(\frac{\sigma}{r_{ij}}\right)^6\left(1-\Delta_{ij}\right) \\
&+ \sum_{Y,Y'}\sum_{i,j}\frac{q_iq_j}{4\pi\epsilon r_{ij}}e^{-\kappa r_{ij}}
\end{aligned}
\tag{C.1}
$$

where $r_{ij}$ is the distance between residues i and j. Each term in Eq. C.1 represents a different set of interactions. The first term accounts for the chain connectivity. The second term accounts for stabilizing interactions between residues that are defined to be in contact based on the crystal structure. The third term represents repulsive long range interactions between non bonded residues. The last term defines non residue specific electrostatic interactions between charged residues. The details of the model can be found in the supplementary information in the ZT work[69].

There are two important energy scales in the system. The first determines the strength of the attractive interactions between the TH and the MT, $\epsilon_h^{TH-MT}$. The second energy scale determines the strength of the docking interactions between the LH and its NL, $\epsilon_h^{LH-NL}$. We chose the values $0.2\frac{KCal}{mol}$ and $0.3\frac{KCal}{mol}$ for $\epsilon_h^{TH-MT}$ and $\epsilon_h^{LH-NL}$ respectively. Simulations using these values produce results that are consistent with the detachment rates of the TH as well as the stall force (Z.Zhang, Y.Goldtzvik, and D.Thirumalai, PNAS, under review).

## C.2 Equations of Motion

We used the equations of motion corresponding to Brownian dynamics in our simulations[103],

$$\mathbf{r}_i(t+h) = \mathbf{r}_i(t) + \frac{h}{\zeta}\mathbf{F}_i + \mathbf{\Gamma}_i(t), \tag{C.2}$$

where $r_i$ is the position of the $i^{th}$ residue, $F_i$ is the force acting on residue i, $\zeta$ is the friction coefficient, and h is the integration step size. We used Stoke's formula, $\zeta = 6\pi\eta a$, with $a = 0.19nm$ to compute the value of the friction coefficient. The random noise, $\Gamma$, was set to have a Gaussian distribution with the property

$$\langle \mathbf{\Gamma}_i(t)\mathbf{\Gamma}_j(t')\rangle = 6\frac{K_B T}{\zeta} h\delta_{ij}\delta_{tt'}. \tag{C.3}$$

where $K_B$ is the Boltzmann constant and T is the temperature.

In order to obtain the correct time scale for the diffusion of the TH, the inclusion of HI is necessary[104]. For the residues of the TH we used the following equation of motion,

$$\mathbf{r}_i(t+h) = \mathbf{r}_i(t) + \sum_j^N \frac{h\mathbf{D}_{ij}}{K_B T}\cdot\mathbf{F}_j + \mathbf{\Gamma}_i(t), \tag{C.4}$$

where $\mathbf{D}_{ij}$ is the Rotne-Prager-Yamakawa diffusion tensor[105,106],

$$\mathbf{D}_{ii} = \frac{(K_B T)}{6\pi\eta a}\mathbf{I}, \tag{C.5}$$

$$\mathbf{D}_{ij} = \frac{K_B T}{8\pi\eta r_{ij}}[(1+\frac{2a^2}{3r_{ij}^2})\mathbf{I} + (1-\frac{2a^2}{r_{ij}^2})\frac{\mathbf{r}_{ij}^T\mathbf{r}_{ij}}{r_{ij}^2}] \qquad r_{ij} \geq 2a, \tag{C.6}$$

$$\mathbf{D}_{ij} = \frac{K_B T}{8\pi\eta r_{ij}}[(\frac{r_{ij}}{2a})(\frac{8}{3}-\frac{3r_{ij}}{4a})\mathbf{I} + \frac{r_{ij}}{4a}\frac{\mathbf{r}_{ij}^T\mathbf{r}_{ij}}{r_{ij}^2}] \qquad r_{ij} < 2a. \tag{C.7}$$

The random forces are correlated in this case and obey,

$$\langle \mathbf{\Gamma}_i(t)\mathbf{\Gamma}_j(t')\rangle = 6\mathbf{D}_{ij}h\delta_{tt'}. \tag{C.8}$$

## C.3 Rotational Constraint of the Coiled Coil

For simulations involving consecutive steps of kinesin, we fixed the orientation of the top two residues of the CC. The constraint was applied in the following manner. We defined $\mathbf{r_{CC}}$ as the vector pointing from the last residue of the CC helix belonging to the TH to its LH counterpart (the top of the CC in Fig. 6.2). We then imposed the following two constraints,

$$\begin{aligned} |\mathbf{r_{CC}}| &= |\mathbf{r_{CC}}|_0 \\ \mathbf{r_{CC}} \cdot \mathbf{r_{CC0}} &= \mathbf{1}. \end{aligned} \tag{C.9}$$

These constraints ensure that the CC at the cargo end does not rotate.

## C.4 Tracking the Position of the TH

In order to monitor the spatial distribution of the TH position in our simulations, we calculated the position of its center of mass and projected it onto a plane $\mathcal{P}$. We define $\mathcal{P}$ as the plane perpendicular to a vector $\hat{\mathbf{n}}$, which in turn was defined as a unit vector, pointing from the center of the MT towards the protofilament on which the motor walks, and perpendicular to the axis of the MT (Fig. C.1). The x axis in $\mathcal{P}$ is the axis of the MT and the y axis is the direction perpendicular to both $\hat{\mathbf{x}}$ and $\hat{\mathbf{n}}$. The origin was set to be the LH center of mass which is a natural choice in order to divide the system into left and right, the x axis being the dividing line.

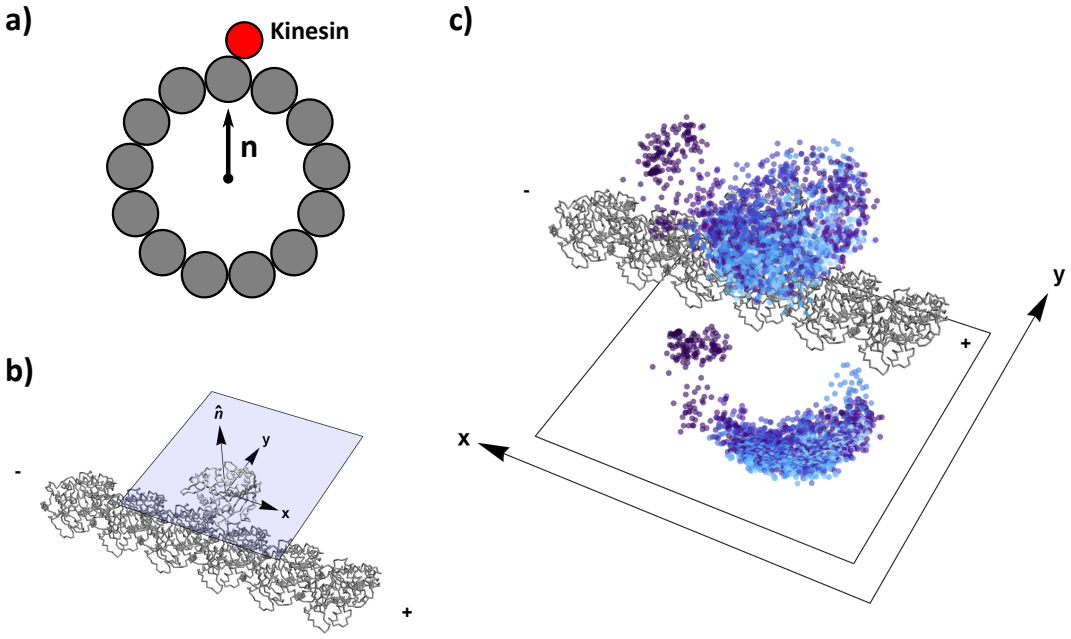


Figure C.1: a) Definition of the vector $\hat{\mathbf{n}}$. The figure shows a cross section of the MT. $\hat{\mathbf{n}}$ points from the center of the MT towards the protofilament on which the kinesin (marked in red) is stepping. b) Definition of the plane $\mathcal{P}$. $\mathcal{P}$ is defined as the plane that is perpendicular to $\hat{\mathbf{n}}$ and that contains the MT axis, x. The axis y is simply the axis perpendicular to both x and $\hat{\mathbf{n}}$. c) An example of the projection for one trajectory. The data points (TH center of mass) in three dimensional space as well as their projection onto $\mathcal{P}$ are shown for a single trajectory.

Positive values along $\hat{\mathbf{y}}$ correspond to the left side while negative values correspond to the right side (Fig. 6.3).

## C.5 Calculations of the Orientation of the TH

While the position of the TH can give insight as to which side of the protofilament is visited more often, it is not a direct measurement of the pathway taken by the TH. A better way is to measure changes in the orientation between the two MDs. We first calculated the end-to-end vector of the TH NL, $\mathbf{r}_{NL}$, and project it onto $\mathcal{P}$. The distribution of these vectors along time can be seen in Fig. 6.5 and 6.11.

In order to quantify the choice of pathway we calculated $\theta$, the angle between the projection of $\mathbf{r}_{NL}$ onto $\mathcal{P}$ and the negative x axis. We allowed the range of $\theta$ to include absolute values larger than 180°. In order to determine the correct value of $\theta$ for a particular trajectory at any point in time, we ensured that the value of $\theta$ as a function of time would not include large jumps. This way, memory of the pathway was retained, allowing us to assess whether the TH reached a point from the left or the right side. It is important to note that due to the noisiness of the data there are data points close to the origin for which it was hard to determine the value of $\theta$ without introducing any jumps. In order to avoid this problem we performed time averaging over 50 frames in order to smooth out the distribution. The distributions shown in Fig. 6.5 and 6.11 are results of such an average.

## C.6 Probability of Symmetric Step

We calculate the probability of kinesin stepping according to the symmetric model when the CC is unconstrained in the following way. We first compute $P_R = 0.69$, the probability of the motor to step from the right side of the LH, directly from the simulations. It follows that the probability of a step from the left side of the LH is $P_L = 0.31$. Assuming that the probability of stepping on either side of the LH is independent from one step to the next due to the long times between the steps, the probability of stepping from the same side twice consecutively is given by,

$$P_S = P_R P_R + P_L P_L \approx 0.57 \tag{C.10}$$

## C.7 NL Tension Calculation

In order to obtain an estimate of the tension along the NL we made the assumption that the NL behaves approximately as a worm-like chain[107]. We calculated $\mathbf{r}_{NL}$, the

vector connecting the first and last residues of the NL, for the LH and projected it onto the x axis, yielding $x_{NL}$. We then used $x_{NL}$ in the force-extension expression,

$$F = \frac{k_B T}{l_P} \left[ \frac{1}{4} \left( 1 - \frac{x_{NL}}{L} \right)^{-2} - \frac{1}{4} + \frac{x_{NL}}{L} \right] \tag{C.11}$$

where $l_P = 0.7nm$ is the persistence length, $L = 0.38\,(N_R - 1)$ is the contour length and $N_R = 13$ is the number of residues in the NL. The Marko-Siggia equation (Eq. C.11) provides near quantitative fits even for small $x_{NL}$ and $l_P$.

# Bibliography


[1] Nobutaka Hirokawa, Yasuko Noda, Yosuke Tanaka, and Shinsuke Niwa. Kinesin superfamily motor proteins and intracellular transport. *Nature reviews Molecular cell biology*, 10(10):682–696, 2009.

[2] John A Hammer and James R Sellers. Walking to work: roles for class v myosins as cargo transporters. *Nature Reviews Molecular Cell Biology*, 13(1):13–26, 2012.

[3] Anthony J Roberts, Takahide Kon, Peter J Knight, Kazuo Sutoh, and Stan A Burgess. Functions and mechanics of dynein motor proteins. *Nature reviews Molecular cell biology*, 14(11):713–726, 2013.

[4] Ronald D Vale. The molecular motor toolbox for intracellular transport. *Cell*, 112(4):467–480, 2003.

[5] Ronald D Vale and Ronald A Milligan. The way things move: looking under the hood of molecular motor proteins. *Science*, 288(5463):88–95, 2000.

[6] Nobutaka Hirokawa. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science*, 279(5350):519–526, 1998.

[7] Amit D Mehta, Ronald S Rock, Matthias Rief, James A Spudich, Mark S Mooseker, and Richard E Cheney. Myosin-v is a processive actin-based motor. *Nature*, 400(6744):590–593, 1999.

[8] Stan A Burgess, Matt L Walker, Hitoshi Sakakibara, Peter J Knight, and Kazuhiro Oiwa. Dynein structure and power stroke. *Nature*, 421(6924):715–718, 2003.

[9] Charles V Sindelar, Mary Jane Budny, Sarah Rice, Nariman Naber, Robert Fletterick, and Roger Cooke. Two conformations in the human kinesin power stroke defined by x-ray crystallography and epr spectroscopy. *Nature Structural & Molecular Biology*, 9(11):844–848, 2002.

[10] Noriyuki Kodera, Daisuke Yamamoto, Ryoki Ishikawa, and Toshio Ando. Video imaging of walking myosin v by high-speed atomic force microscopy. *Nature*, 468(7320):72–76, 2010.

[11] Ahmet Yildiz, Joseph N Forkey, Sean A McKinney, Taekjip Ha, Yale E Goldman, and Paul R Selvin. Myosin v walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *science*, 300(5628):2061–2065, 2003.

[12] Steven M Block, Lawrence SB Goldstein, and Bruce J Schnapp. Bead movement by single kinesin molecules studied with optical tweezers. *Nature*, 348:348âĂŞ–352, 1990.

[13] Mark A DeWitt, Amy Y Chang, Peter A Combs, and Ahmet Yildiz. Cytoplasmic dynein moves through uncoordinated stepping of the aaa+ ring domains. *Science*, 335(6065):221–225, 2012.

[14] Weihong Qiu, Nathan D Derr, Brian S Goodman, Elizabeth Villa, David Wu, William Shih, and Samara L Reck-Peterson. Dynein achieves processive motion using both stochastic and coordinated stepping. *Nature structural & molecular biology*, 19(2):193–200, 2012.

[15] F Jon Kull, Elena P Sablin, Rebecca Lau, Robert J Fletterick, and Ronald D

Vale. Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature*, 380(6574):550, 1996.

[16] Zhiguo Shang, Kaifeng Zhou, Chen Xu, Roseann Csencsits, Jared C Cochran, and Charles V Sindelar. High-resolution structures of kinesin on microtubules provide a basis for nucleotide-gated force-generation. *Elife*, 3:e04686, 2014.

[17] Andrew P Carter, Carol Cho, Lan Jin, and Ronald D Vale. Crystal structure of the dynein motor domain. *Science*, 331(6021):1159–1165, 2011.

[18] Takahide Kon, Takuji Oyama, Rieko Shimo-Kon, Kenji Imamula, Tomohiro Shima, Kazuo Sutoh, and Genji Kurisu. The 2.8 Å crystal structure of the dynein motor domain. *Nature*, 484(7394):345–350, 2012.

[19] Helgo Schmidt, Ruta Zalyte, Linas Urnavicius, and Andrew P Carter. Structure of human cytoplasmic dynein-2 primed for its power stroke. *Nature*, 518(7539):435–438, 2015.

[20] Anthony J Roberts, Naoki Numata, Matt L Walker, Yusuke S Kato, Bara Malkova, Takahide Kon, Reiko Ohkura, Fumio Arisaka, Peter J Knight, Kazuo Sutoh, et al. Aaa+ ring and linker swing mechanism in the dynein motor. *Cell*, 136(3):485–495, 2009.

[21] William B Redwine, Rogelio Hernández-López, Sirui Zou, Julie Huang, Samara L Reck-Peterson, and Andres E Leschziner. Structural basis for microtubule binding and release by dynein. *Science*, 337(6101):1532–1536, 2012.

[22] Trina A Schroer, Eric R Steuer, and Michael P Sheetz. Cytoplasmic dynein is a minus end-directed motor for membranous organelles. *Cell*, 56(6):937–946, 1989.

[23] Keith E Summers and IR Gibbons. Adenosine triphosphate-induced sliding of tubules in trypsin-treated flagella of sea-urchin sperm. *Proceedings of the National Academy of Sciences*, 68(12):3092–3096, 1971.

[24] Gregory J Pazour, Bethany L Dickert, and George B Witman. The dhc1b (dhc2) isoform of cytoplasmic dynein is required for flagellar assembly. *The Journal of cell biology*, 144(3):473–481, 1999.

[25] Mary E Porter, Raqual Bower, Julie A Knott, Pamela Byrd, and William Dentler. Cytoplasmic dynein heavy chain 1b is required for flagellar assembly in chlamydomonas. *Molecular biology of the cell*, 10(3):693–712, 1999.

[26] Owen J Driskell, Aleksandr Mironov, Victoria J Allan, and Philip G Woodman. Dynein is required for receptor sorting and the morphogenesis of early endosomes. *Nature Cell Biology*, 9(1):113–120, 2007.

[27] Ingrid Jordens, Mar Fernandez-Borja, Marije Marsman, Simone Dusseljee, Lennert Janssen, Jero Calafat, Hans Janssen, Richard Wubbolts, and Jacques Neefjes. The rab7 effector protein rilp controls lysosomal transport by inducing the recruitment of dynein-dynactin motors. *Current Biology*, 11(21):1680–1685, 2001.

[28] Aaron D Pilling, Dai Horiuchi, Curtis M Lively, and William M Saxton. Kinesin-1 and dynein are the primary motors for fast transport of mitochondria in drosophila motor axons. *Molecular biology of the cell*, 17(4):2057–2068, 2006.

[29] Aaron Young, Jason B Dictenberg, Aruna Purohit, Richard Tuft, and Stephen J Doxsey. Cytoplasmic dynein-mediated assembly of pericentrin and $\gamma$ tubulin onto centrosomes. *Molecular Biology of the Cell*, 11(6):2047–2056, 2000.

[30] Jennifer M Harrell, Patrick JM Murphy, Yoshihiro Morishima, Haifeng Chen, John F Mansfield, Mario D Galigniana, and William B Pratt. Evidence for glucocorticoid receptor transport on microtubules by dynein. *Journal of Biological Chemistry*, 279(52):54647–54654, 2004.

[31] Liedewij Laan, Nenad Pavin, Julien Husson, Guillaume Romet-Lemonne, Martijn Van Duijn, Magdalena Preciado López, Ronald D Vale, Frank Jülicher, Samara L Reck-Peterson, and Marileen Dogterom. Cortical dynein controls microtubule dynamics to generate pulling forces that position microtubule asters. *Cell*, 148(3):502–514, 2012.

[32] Francis J McNally. Mechanisms of spindle positioning. *The Journal of cell biology*, 200(2):131–140, 2013.

[33] Comert Kural, Hwajin Kim, Sheyum Syed, Gohta Goshima, Vladimir I Gelfand, and Paul R Selvin. Kinesin and dynein move a peroxisome in vivo: a tug-of-war or coordinated movement? *Science*, 308(5727):1469–1472, 2005.

[34] Melanie JI Müller, Stefan Klumpp, and Reinhard Lipowsky. Tug-of-war as a cooperative mechanism for bidirectional cargo transport by molecular motors. *Proceedings of the National Academy of Sciences*, 105(12):4609–4614, 2008.

[35] Virupakshi Soppina, Arpan Kumar Rai, Avin Jayesh Ramaiya, Pradeep Barak, and Roop Mallik. Tug-of-war between dissimilar teams of microtubule motors regulates transport and fission of endosomes. *Proceedings of the National Academy of Sciences*, 106(46):19381–19386, 2009.

[36] Adam G Hendricks, Eran Perlson, Jennifer L Ross, Harry W Schroeder, Mariko Tokito, and Erika LF Holzbaur. Motor coordination via a tug-of-war mechanism drives bidirectional vesicle transport. *Current Biology*, 20(8):697–702, 2010.

[37] William O Hancock. Bidirectional cargo transport: Moving beyond tug-of-war. *Nature reviews. Molecular cell biology*, 15(9):615, 2014.

[38] Andrew F Neuwald, L Aravind, John L Spouge, and Eugene V Koonin. Aaa+: A class of chaperone-like atpases associated with the assembly, operation, and disassembly of protein complexes. *Genome research*, 9(1):27–43, 1999.

[39] Samara L Reck-Peterson, Ahmet Yildiz, Andrew P Carter, Arne Gennerich, Nan Zhang, and Ronald D Vale. Single-molecule analysis of dynein processivity and stepping behavior. *Cell*, 126(2):335–348, 2006.

[40] Takahide Kon, Masaya Nishiura, Reiko Ohkura, Yoko Y Toyoshima, and Kazuo Sutoh. Distinct functions of nucleotide-binding/hydrolysis sites in the four aaa modules of cytoplasmic dynein. *Biochemistry*, 43(35):11266–11274, 2004.

[41] Kenji Imamula, Takahide Kon, Reiko Ohkura, and Kazuo Sutoh. The coordination of cyclic microtubule association/dissociation and tail swing of cytoplasmic dynein. *Proceedings of the National Academy of Sciences*, 104(41):16134–16139, 2007.

[42] Helgo Schmidt, Emma S Gleave, and Andrew P Carter. Insights into dynein motor domain function from a 3.3-å crystal structure. *Nature structural & molecular biology*, 19(5):492–497, 2012.

[43] Gira Bhabha, Hui-Chun Cheng, Nan Zhang, Arne Moeller, Maofu Liao, Jeffrey A Speir, Yifan Cheng, and Ronald D Vale. Allosteric communication in the dynein motor domain. *Cell*, 159(4):857–868, 2014.

[44] Mark A DeWitt, Caroline A Cypranowska, Frank B Cleary, Vladislav Belyy, and Ahmet Yildiz. The aaa3 domain of cytoplasmic dynein acts as a switch to

facilitate microtubule release. *Nature structural & molecular biology*, 22(1):73–80, 2015.

[45] Matthew P Nicholas, Florian Berger, Lu Rao, Sibylle Brenner, Carol Cho, and Arne Gennerich. Cytoplasmic dynein regulates its attachment to microtubules via nucleotide state-switched mechanosensing at multiple aaa domains. *Proceedings of the National Academy of Sciences*, 112(20):6371–6376, 2015.

[46] Takahide Kon, Toshifumi Mogami, Reiko Ohkura, Masaya Nishiura, and Kazuo Sutoh. Atp hydrolysis cycle–dependent tail motions in cytoplasmic dynein. *Nature structural & molecular biology*, 12(6):513–519, 2005.

[47] Frank B Cleary, Mark A Dewitt, Thomas Bilyard, Zaw Min Htet, Vladislav Belyy, Danna D Chan, Amy Y Chang, and Ahmet Yildiz. Tension on the linker gates the atp-dependent release of dynein from microtubules. *Nature communications*, 5, 2014.

[48] Peter Höök, Atsushi Mikami, Beth Shafer, Brian T Chait, Steven S Rosenfeld, and Richard B Vallee. Long range allosteric control of cytoplasmic dynein atpase activity by the stalk and c-terminal domains. *Journal of Biological Chemistry*, 280(38):33045–33054, 2005.

[49] Toshifumi Mogami, Takahide Kon, Kohji Ito, and Kazuo Sutoh. Kinetic characterization of tail swing steps in the atpase cycle of dictyostelium cytoplasmic dynein. *Journal of Biological Chemistry*, 282(30):21639–21644, 2007.

[50] Takahide Kon, Kenji Imamula, Anthony J Roberts, Reiko Ohkura, Peter J Knight, IR Gibbons, Stan A Burgess, and Kazuo Sutoh. Helix sliding in the stalk coiled coil of dynein couples atpase and microtubule binding. *Nature structural & molecular biology*, 16(3):325–333, 2009.

[51] IR Gibbons, Joan E Garbarino, Carol E Tan, Samara L Reck-Peterson, Ronald D Vale, and Andrew P Carter. The affinity of the dynein microtubule-binding domain is modulated by the conformation of its coiled-coil stalk. *Journal of Biological Chemistry*, 280(25):23960–23965, 2005.

[52] Andrew P Carter, Joan E Garbarino, Elizabeth M Wilson-Kubalek, Wesley E Shipley, Carol Cho, Ronald A Milligan, Ronald D Vale, and IR Gibbons. Structure and functional role of dynein's microtubule-binding domain. *Science*, 322(5908):1691–1695, 2008.

[53] Yosuke Nishikawa, Momoko Inatomi, Haruka Iwasaki, and Genji Kurisu. Structural change in the dynein stalk region associated with two different affinities for the microtubule. *Journal of Molecular Biology*, 428(9):1886–1896, 2016.

[54] Erika LF Holzbaur and Kenneth A Johnson. Microtubules accelerate adp release by dynein. *Biochemistry*, 28(17):7010–7016, 1989.

[55] Tadashi Ando and Jeffrey Skolnick. Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. *Proceedings of the National Academy of Sciences*, 107(43):18457–18462, 2010.

[56] David D Hackney. The tethered motor domain of a kinesin-microtubule complex catalyzes reversible synthesis of bound atp. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51):18338–18343, 2005.

[57] Claudia Veigel, Fei Wang, Marc L Bartoo, James R Sellers, and Justin E Molloy. The gated gait of the processive molecular motor, myosin v. *Nature cell biology*, 4(1):59–65, 2002.

[58] Sotaro Uemura, Kenji Kawaguchi, Junichiro Yajima, Masaki Edamatsu, Yoko Yano Toyoshima, and Shin'ichi Ishiwata. Kinesin–microtubule binding depends on both nucleotide state and loading direction. *Proceedings of the National Academy of Sciences*, 99(9):5977–5981, 2002.

[59] Stanislas Leibler and David A Huse. Porters versus rowers: a unified stochastic model of motor proteins. *The Journal of cell biology*, 121(6):1357–1368, 1993.

[60] Michael E Fisher and Anatoly B Kolomeisky. The force exerted by a molecular motor. *Proceedings of the National Academy of Sciences*, 96(12):6597–6602, 1999.

[61] Anatoly B Kolomeisky and Michael E Fisher. A simple kinetic model describes the processivity of myosin-v. *Biophysical journal*, 84(3):1642–1650, 2003.

[62] Denis Tsygankov, Adrian WR Serohijos, Nikolay V Dokholyan, and Timothy C Elston. Kinetic models for the coordinated stepping of cytoplasmic dynein. *The Journal of chemical physics*, 130(2):025101, 2009.

[63] Ahmet Yildiz, Michio Tomishige, Ronald D Vale, and Paul R Selvin. Kinesin walks hand-over-hand. *Science*, 303(5658):676–678, 2004.

[64] Sanghamitra Ray, Edgar Meyhöfer, Ronald A Milligan, and Jonathon Howard. Kinesin follows the microtubule's protofilament axis. *The Journal of Cell Biology*, 121(5):1083–1093, 1993.

[65] Denis Tsygankov, Adrian WR Serohijos, Nikolay V Dokholyan, and Timothy C Elston. A physical model reveals the mechanochemistry responsible for dynein's processive motion. *Biophysical journal*, 101(1):144–150, 2011.

[66] Andreja Šarlah and Andrej Vilfan. The winch model can explain both coordi-

nated and uncoordinated stepping of cytoplasmic dynein. *Biophysical journal*, 107(3):662–671, 2014.

[67] Michael Hinczewski, Riina Tehver, and D Thirumalai. Design principles governing the motility of myosin v. *Proceedings of the National Academy of Sciences*, 110(43):E4059–E4068, 2013.

[68] Riina Tehver and D Thirumalai. Rigor to post-rigor transition in myosin v: link between the dynamics and the supporting architecture. *Structure*, 18(4):471–481, 2010.

[69] Zhechun Zhang and D Thirumalai. Dissecting the kinematics of the kinesin step. *Structure*, 20(4):628–640, 2012.

[70] Carol Cho, Samara L Reck-Peterson, and Ronald D Vale. Regulatory atpase sites of cytoplasmic dynein affect processivity and force generation. *Journal of Biological Chemistry*, 283(38):25839–25845, 2008.

[71] Hironori Ueno, Takuo Yasunaga, Chikako Shingyoji, and Keiko Hirose. Dynein pulls microtubules without rotating its stalk. *Proceedings of the National Academy of Sciences*, 105(50):19702–19707, 2008.

[72] Masaomi Nangaku, Reiko Sato-Yoshitake, Yasushi Okada, Yasuko Noda, Reiko Takemura, Hiroto Yamazaki, and Nobutaka Hirokawa. Kif1b, a novel microtubule plus end-directed monomeric motor protein for transport of mitochondria. *Cell*, 79(7):1209–1220, 1994.

[73] Sen Takeda, Hiroto Yamazaki, Dae-Hyun Seog, Yoshimitsu Kanai, Sumio Terada, and Nobutaka Hirokawa. Kinesin superfamily protein 3 (kif3) motor transports fodrin-associating vesicles important for neurite building. *The Journal of cell biology*, 148(6):1255–1266, 2000.

[74] KEITH G KozMINSKI, Karl A Johnson, Paul Forscher, and Joel L Rosenbaum. A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proceedings of the National Academy of Sciences*, 90(12):5519–5523, 1993.

[75] Tobias Stauber, Jeremy C Simpson, Rainer Pepperkok, and Isabelle Vernos. A role for kinesin-2 in copi-dependent recycling between the er and the golgi complex. *Current biology*, 16(22):2245–2251, 2006.

[76] Takao Nakata and Nobutaka Hirokawa. Point mutation of adenosine triphosphate-binding motif generated rigor kinesin that selectively blocks anterograde lysosome membrane transport. *The Journal of cell biology*, 131(4):1039–1053, 1995.

[77] Eustratios Bananis, John W Murray, Richard J Stockert, Peter Satir, and Allan W Wolkoff. Microtubule and motor-dependent endocytic vesicle sorting in vitro. *The Journal of cell biology*, 151(1):179–186, 2000.

[78] Stefan Sack, Jens Müller, Alexander Marx, Manfred Thormählen, Eva-Maria Mandelkow, Scott T Brady, and Eckhard Mandelkow. X-ray structure of motor and neck domains from rat brain kinesin. *Biochemistry*, 36(51):16155–16165, 1997.

[79] Sarah Rice, Abel W Lin, Daniel Safer, Cynthia L Hart, Nariman Naber, Bridget O Carragher, Shane M Cain, Elena Pechatnikova, Elizabeth M Wilson-Kubalek, Michael Whittaker, et al. A structural change in the kinesin motor protein that drives motility. *Nature*, 402(6763):778–784, 1999.

[80] Michio Tomishige, Nico Stuurman, and Ronald D Vale. Single-molecule observations of neck linker conformational changes in the kinesin motor protein. *Nature structural & molecular biology*, 13(10):887–894, 2006.

[81] Nicholas R Guydosh and Steven M Block. Backsteps induced by nucleotide analogs suggest the front head of kinesin is gated by strain. *Proceedings of the National Academy of Sciences*, 103(21):8054–8059, 2006.

[82] Ahmet Yildiz, Michio Tomishige, Arne Gennerich, and Ronald D Vale. Intramolecular strain coordinates kinesin stepping behavior along microtubules. *Cell*, 134(6):1030–1041, 2008.

[83] Merve Yusra Dogan, Sinan Can, Frank B Cleary, Vedud Purde, and Ahmet Yildiz. Kinesin's front head is gated by the backward orientation of its neck linker. *Cell reports*, 10(12):1967–1973, 2015.

[84] Charles L Asbury, Adrian N Fehr, and Steven M Block. Kinesin moves by an asymmetric hand-over-hand mechanism. *Science*, 302(5653):2130–2134, 2003.

[85] David D Hackney. Evidence for alternating head catalysis by kinesin during microtubule-stimulated atp hydrolysis. *Proceedings of the National Academy of Sciences*, 91(15):6865–6869, 1994.

[86] Nick J Carter and RA Cross. Mechanics of the kinesin step. *Nature*, 435(7040):308–312, 2005.

[87] Braulio Gutiérrez-Medina, Adrian N Fehr, and Steven M Block. Direct measurements of kinesin torsional properties reveal flexible domains and occasional stalk reversals during stepping. *Proceedings of the National Academy of Sciences*, 106(40):17007–17012, 2009.

[88] Hiroshi Isojima, Ryota Iino, Yamato Niitani, Hiroyuki Noji, and Michio Tomishige. Direct observation of intermediate states during the stepping motion of kinesin-1. *Nature chemical biology*, 2016.

[89] Changbong Hyeon and José N Onuchic. Mechanical control of the directional stepping dynamics of the kinesin motor. *Proceedings of the National Academy of Sciences*, 104(44):17382–17387, 2007.

[90] Wei Hua, Johnson Chung, and Jeff Gelles. Distinguishing inchworm and hand-over-hand processive kinesin movement by neck rotation measurements. *Science*, 295(5556):844–848, 2002.

[91] Changbong Hyeon and José N Onuchic. Internal strain regulates the nucleotide binding site of the kinesin leading head. *Proceedings of the National Academy of Sciences*, 104(7):2175–2180, 2007.

[92] Teppei Mori, Ronald D Vale, and Michio Tomishige. How kinesin waits between steps. *Nature*, 450(7168):750–754, 2007.

[93] Paul C Whitford, Peter Geggier, Roger B Altman, Scott C Blanchard, José N Onuchic, and Karissa Y Sanbonmatsu. Accommodation of aminoacyl-trna into the ribosome involves reversible excursions along multiple pathways. *Rna*, 16(6):1196–1204, 2010.

[94] Changbong Hyeon and D Thirumalai. Capturing the essence of folding and functions of biomolecules using coarse-grained models. *Nature communications*, 2:487, 2011.

[95] Changbong Hyeon, George H Lorimer, and D Thirumalai. Dynamics of allosteric transitions in groel. *Proceedings of the National Academy of Sciences*, 103(50):18939–18944, 2006.

[96] Changbong Hyeon, Ruxandra I Dima, and D Thirumalai. Pathways and kinetic barriers in mechanical unfolding and refolding of rna and proteins. *Structure*, 14(11):1633–1645, 2006.

[97] Jie Chen, Ruxandra I Dima, and D Thirumalai. Allosteric communication in dihydrofolate reductase: signaling network and pathways for closed to occluded transition and back. *Journal of molecular biology*, 374(1):250–266, 2007.

[98] Seiichi Uchimura, Takashi Fujii, Hiroko Takazaki, Rie Ayukawa, Yosuke Nishikawa, Itsushi Minoura, You Hachikubo, Genji Kurisu, Kazuo Sutoh, Takahide Kon, et al. A flipped ion pair at the dynein–microtubule interface is critical for dynein motility and atpase activation. *J Cell Biol*, 208(2):211–222, 2015.

[99] Junghyun Choi, Hahnbeom Park, and Chaok Seok. How does a registry change in dynein's coiled-coil stalk drive binding of dynein to microtubules? *Biochemistry*, 50(35):7629–7636, 2011.

[100] Joshua W Shaevitz, Steven M Block, and Mark J Schnitzer. Statistical kinetics of macromolecular dynamics. *Biophysical journal*, 89(4):2277–2285, 2005.

[101] Wolfram Research, Inc. Mathematica, Version 10.0. Champaign, IL, 2014.

[102] Elijah Roberts, John Eargle, Dan Wright, and Zaida Luthey-Schulten. Multiseq: unifying sequence and structure data for evolutionary analysis. *BMC bioinformatics*, 7(1):382, 2006.

[103] Donald L Ermak and JA McCammon. Brownian dynamics with hydrodynamic interactions. *The Journal of chemical physics*, 69(4):1352–1360, 1978.

[104] Yonathan Goldtzvik, Zhechun Zhang, and D Thirumalai. Importance of hydrodynamic interactions in the stepping kinetics of kinesin. *The Journal of Physical Chemistry B*, 120(8):2071–2075, 2016.

[105] Jens Rotne and Stephen Prager. Variational treatment of hydrodynamic interaction in polymers. *The Journal of Chemical Physics*, 50(11):4831–4837, 1969.

[106] Hiromi Yamakawa. Transport properties of polymer chains in dilute solution: hydrodynamic interaction. *The Journal of Chemical Physics*, 53(1):436–443, 1970.

[107] John F Marko and Eric D Siggia. Stretching dna. *Macromolecules*, 28(26):8759–8770, 1995.